DOCUMENT RESUME

ED 063 807	24	EM 010 165

AUTHOR Spencer, Rosemary Y.; Briggs, Leslie J.

TITLE Application of a Learning Hierarchy to Sequence an Instructional Program, and Comparison of this Program with Reverse and Random Sequences. Final Report.

INSTITUTION Florida State Univ., Tallahassee. Dept. of Educational Research.

SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C.

BUREAU NO BR-1-D-007

PUB DATE Feb 72

GRANT OEG-4-71-0071

NOTE 151p.

EDRS PRICE MF-$0.65 HC-$6.58

DESCRIPTORS *Algebra; *Intermode Differences; *Learning Theories; Mathematics Instruction; *Programed Instruction; Secondary School Mathematics; *Task Analysis; Transfer of Training

IDENTIFIERS *Gagne (R M)

ABSTRACT

Two studies attempted to determine whether a programed instruction booklet on a task in algebra would be more effective if the units in the program were presented in a sequence suggested by a hierarchical analysis of the task than if units were presented in a reverse or a random sequence. In the first study, a program was made up, based on the Princeton Algebra Program, in which blocks of frames made up units corresponding to the competencies in the hierarchy developed by Gagne and Paradise. Three versions of the program--forward order, reverse order, random order--were used with groups of eighth-graders in a school with a large proportion of disadvantaged students. The results yielded many trends in differences among the three versions of the program, but few were statistically significant. The second study was a replication using older students from more advantaged background and only the forward and random order programs. The forward version of the program resulted in better performance for the older and more advantaged children than for the younger ones in the first study. The highest retention scores however came from younger, less advantaged subjects who had the random sequence in the first study. These results and other significant findings are discussed in some detail. (JY)

BR 1-D-007
PA 24

Final Report

Project No. 1-D-007
Grant No. OEG-4-71-0071

Rosemary Y. Spencer and Leslie J. Briggs
Department of Educational Research
College of Education
Florida State University
Tallahassee, Florida 32306

Application Of A Learning Hierarchy To Sequence An Instructional Program, And Comparison Of This Program With Reverse And Random Sequences

February, 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development
(Regional Research Program)

APPLICATION OF A LEARNING HIERARCHY TO SEQUENCE AN INSTRUCTIONAL PROGRAM, AND COMPARISON OF THIS PROGRAM WITH REVERSE AND RANDOM SEQUENCES

Rosemary Y. Spencer and Leslie J. Briggs[1]
The Florida State University

Abstract

This was a follow-on study (Experiment II) to investigate further whether a programmed instruction booklet would be more effective if the units of the program were presented in a sequence implied by a "learning hierarchy," as compared to a random sequence.

The program was sequenced by classifying individual program frames to correspond with competencies in the hierarchy, thus forming instructional units which corresponded to the hierarchical competencies. The program so sequenced was subjected to three revisions to establish teaching effectiveness prior to Experiment I.

In Experiment I, sequence effects by I.Q. levels were studied, comparing hierarchy (forward) sequence with reverse and random sequences, using eighth-grade pupils in a school characterized as having a large proportion of underpriviledged children.

In the present study, Experiment II, a larger number of older pupils from more priviledged homes were used as Ss, and the study was replicated only for the forward and the random sequences.

The results of Experiment II were that the forward sequence group (as compared to the random sequence group) took less time to complete the program ($p < .01$) and had superior performance on the posttest ($p < .05$). Also, low ability Ss in the forward group performed consistently better than did the low-ability Ss in the random group; low ability Ss, particularly, benefitted from the hierarchical sequence.

[1]The authors express appreciation to Dr. F. J. King, who served as statistical consultant for the study. It was he who called the authors' attention to the relevance of path analysis to the problem, as discussed under Recommendations for Future Research.

Final Report

Project No. 1-D-007
Grant No. OEG-4-71-0071

Application of a Learning Hierarchy to Sequence An Instructional Program, and Comparison of This Program With Reverse and Random Sequences

Volume II of II Volumes

Rosemary Y. Spencer and Leslie J. Briggs
The Florida State University
Tallahassee, Florida

February, 1972

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

APPLICATION OF A LEARNING HIERARCHY TO SEQUENCE AN INSTRUCTIONAL PROGRAM, AND COMPARISON OF THIS PROGRAM WITH REVERSE AND RANDOM SEQUENCES

Problem

In Volume I of this Final Report, the background of the problem was presented in great detail, and the results of an experiment (Experiment I) were presented.

In this Volume II, the results of a follow-on study (Experiment II) are reported.

The reader is referred to Volume I for a comprehensive view of the prior work. Only an extremely abbreviated summary is given here.

The purpose of Experiment I was to determine whether a programmed instruction booklet on a task in algebra would be more effective if the units in the program were presented in a sequence suggested by a hierarchical analysis of the task than if units were presented in a reverse or a random sequence.

A hierarchy for a task which was previously developed by Gagné and Paradise (1961) was taken as a starting point in the work leading up to Experiment I. The total task used by those investigators was that of Solving Equations. For Experiment I, only one branch of that hierarchy was used; this was the branch referred to as task I,1, in Volume I of this report. Thus the total "task" used in Experiment I was essentially a "subtask" from the study by Gagné and Paradise.

The subordinate competencies listed by Gagné and Paradise for task I,1 were taken as the reference point for developing a new programmed instruction booklet. Frames from the Princeton Algebra Program, used by Gagné and Paradise, were sorted out so as to identify which subordinate competency in the hierarchy each frame supported. Then new frames were written, or old ones revised, until it was believed there were adequate frames for each competency to make up an effective teaching "unit." This resulted in a "program" in which blocks of frames made up "units" corresponding to the competencies in the hierarchy developed by Gagné and Paradise.

Using the resulting new program (greatly changed from the original program), empirical tryouts and revisions were conducted to make the program as effective as possible. This method of program development and improvement resulted in the "forward" version of the program --the version in which the units were sequenced in accordance with the hierarchy.

Then it was a simple matter to rearrange the units (but keeping the frame sequence intact within units) to make up the "reverse" and the "random" versions of the program.

Experiment I consisted of a comparison of results from use of the three versions of the program with groups of eighth-grade pupils in a school characterized as having a large proportion of underpriviledged children.

The results from Experiment I yielded many trends in differences among the three versions of the program, but there were few statistically significant differences. It was believed that different results might be found if larger numbers of pupils were used in a second experiment, and if these pupils came from more advantaged homes. It was also thought that since algebra is usually studied later than the eighth grade, a second experiment with older pupils (but only those not previously having studied algebra) might yield results more typical of pupils who study algebra.

In summary, it was believed there were several considerations justifying a repetition of the study. The importance of the problem is a major justification. Also, a larger sample from a population different from that employed in Experiment I would provide a better basis both for reaching a summative evaluation of the practical effectiveness of the forward version of the program, and for reaching conclusions concerning hypotheses earlier tested in Experiment I.

To maximize the number of learners in each group, for Experiment II, it was decided to replicate only the "forward" and the "random" sequences from Experiment I. Another reason for this was that there were few differences in the results between "reverse" and "random" sequences in Experiment I. Finally, it was thought that conventional, "loosely structured" materials more closely resemble the "random" than the "reverse" program, thus making the "random" sequence more akin than would be the "reverse" sequence to conventional materials that are not sequenced according to a hierarchy.

Experiment II

In Experiment II, results for the learning program having instructional units sequenced in the order implied by task I,1 of the hierarchy (forward sequence) were compared with results for a random sequence of unit presentation. The random sequence had been determined for Experiment I by randomly drawing slips of paper numbered 1 to 10 from a box. The sequence so determined, coded for competencies listed in Volume I, was: 8,4,9,5,10,7,6,1,3,2. The sequence of frames within units, as in Experiment I, was left constant, while varying the sequence of presentation of the intact units (groups of frames). Also, as in Experiment I, treatment by levels analyses of variance were used to evaluate treatment effects (sequence variations) upon performance, as well as the interaction of mental ability and sequence upon performance.

The posttest performance mean for the forward sequence group was considered as a second summative evaluation of the program.

Subjects

The programs and tests from Experiment I were administered to 285 Ss in the ninth through the twelfth grades in 10 beginning algebra classrooms in Tallahassee, Florida. The materials employed are described in detail in Volume I of this report and more briefly here. As in Experiment I, results were not used for Ss having a pretest score of eight or above. Also excluded from the results were Ss of I.Q. 80 or below, and Ss for whom no I.Q. was available. These restrictions resulted in use of data for 175 of the total 285 Ss in the intact classrooms in which Experiment II was conducted.

Measures

In Experiment II, as in Experiment I, a competency test, a pretest and identical posttest, a retention test, and an attitude scale were used.

The competency test consisted of the representative test item used by Gagné and Paradise in defining each of the 10 subordinate competencies of task I,1 of the hierarchy. The competency test items were scattered throughout the program as the last frame of each instructional unit; each such test frame was identified as a "Test Question," and no feedback was given on the test questions in Experiment II, while feedback was given in Experiment I.

The 10-item pretest and identical posttest used to evaluate performance on final task I,1 was constructed by writing test items varying in complexity, but all similar to the representative item used by Gagné and Paradise to define task I,1.

The retention test was considered a parallel form of the posttest, since the 10 items on the two were matched in complexity.

The attitude scale was used for the purpose of estimating Ss attitude toward the program. It was expected that the forward sequence group would have a more favorable attitude toward the program than would the random sequence group. The attitude scale was developed according to recommended techniques, and was entitled "Learning Program Questionnaire."

Procedures

Administration of the learning program, tests, and attitude scale was carried out in each classroom by the teacher and a proctor.

The pretest was administered the first Friday school was in session in September, 1971, in the 10 classes to take part in the study. Seventeen Ss in the 10 classrooms scored eight or above on the pretest. Since these 17 Ss met or surpassed the design criterion for the instruction, their results were not used for the experiment.

Individuals were randomly assigned within stratified I.Q. level to the two sequence presentations; Ss names were arranged by score on the California Test of Mental Maturity, from low to high. Then, dividing the group by thirds, three mental ability levels were formed whose I.Q. ranges were: high, 122-142; average, 108-121; and low, 80-107. A coin was flipped assigning the first name on the list to the random sequence, the next name the forward sequence, next the random sequence, and so on down the list of names.

On the Monday following the pretest administration, each S received a learning program coded to his experimental treatment, with his name written on it. Two pages of directions for taking the program on a completely self-instructional basis were read aloud to the Ss, as they read them to themselves at their desks. In addition, the Ss were informed:

> "All of you have the same learning program, but the program has been put together in two different ways so that the order of instruction varies. A list of your names has been used to randomly assign you to one of two types of instructional sequence."

As Ss finished the learning program, the date and time were noted on the front of their program booklets. Then, after Ss completed the additude scale, they received a posttest.

Ss finished the program during that school week. The retention test was administered three weeks later.

Results of Experiment II

A 2 x 3 analysis of variance was used to study the two sequence variations by three I.Q. levels. In Table 1, the analysis of variance F ratios support the conclusion that the forward group took significantly less time to complete the program ($F=6.99$, df 1/169, $p<.01$) and performed better on the posttest ($F=4.11$, df 1/167, $p<.05$). There were no significant differences between the forward and random sequence groups for the variables: program errors; competency test score; attitude measure; or retention.

Though the Experiment I finding of significantly fewer program errors for the forward group was not found for this Experiment II, a study of the means in Table 2 reveals the trend in favor of the forward group.

Interaction between sequence presentation and mental ability level on attitude toward the program (see Table 1) was significant ($F=3.42$, df 2/169, $p<.05$). No other significant interaction effects were found for sequence and mental ability upon performance.

Perhaps the most important result of Experiment II, which supports the findings of Experiment I, was that low ability Ss in the forward group performed consistently better than did low ability Ss in the random group. According to Table 2, low ability Ss who received the forward sequence: (a) took less time to complete the program; (b) had fewer program errors; (c) had a higher competency test score; (d) indicated a more favorable attitude toward the program; (e) performed better on the posttest; and (f) had superior retention. These trends also appear for Experiment I, except for attitude scores. Thus the means for low I.Q. Ss favor forward over random in 11 of 12 comparisons. On the other hand, the Ss of high ability and of average ability in the forward group did not perform consistently better than their counterparts in the random sequence group.

Tables 3 and 4 present the intercorrelations among I.Q. score, pretest score, and the six dependent variables under study, for the forward group (Table 3) and random group (Table 4). The forward group I.Q. and posttest score correlation ($r=.32$), and the random group I.Q. and posttest score correlation ($r=.58$) were significantly different ($p<.05$). Thus ability played a less important part in posttest achievement for the forward sequence than it did for the random sequence.

This result implies an interaction of I.Q. and sequence treatment upon posttest score, which is also suggested in study of the trends among posttest means by I.Q. level, even though the analysis of variance does not show significant interaction (Table 1). Thus the correlations, representing scores on continuous variables, suggest an

interaction not confirmed by the grouped data used for the analysis of variance. This implied interaction suggests that the "better" the instructional program, the less effect I.Q. has upon posttest success. Said differently, careful sequencing can bring posttest success to low ability Ss who might not have succeeded without such careful sequencing.

The relationship between posttest score and retention test score for the forward group was r=.58, and for the random group, r=.74. These positive correlations within treatments are different in apparent meaning from the rank order of means, among treatments, for posttest and retention test.

The results of the forward sequence group (means for the competency test and the posttest), taken as a second summative evaluation of the program, were: competency test 71/73 (71% of Ss scored 73% or higher) and posttest 60/60.

Discussion

The present study can be distinguished from two earlier groups of experiments relating to sequencing of instruction and/or to learning hierarchies. The research of Gagné and his associates was involved primarily with testing the predictions of transfer among adjacent pairs of subordinate learning sets (competencies) of a learning hierarchy supporting a final task. These studies were not concerned with relating a total instructional sequence to the overall arrangement of competencies of the hierarchy.

In contrast to the studies by Gagné, in the experiments reported here, a program was sequenced according to a learning hierarchy, and then the entire instructional program so sequenced was compared with a random program version (and also a reverse sequence for Experiment I). Thus the present studies are concerned with the overall applicability of a hierarchy for the purpose of arranging the sequencing of instruction, not for the purpose of studying transfer among single pairs of competencies.

The majority of what are termed sequencing studies in the literature in programmed instruction are experiments comparing logical versus scrambled frame (rather than unit) sequences. They do not attempt to apply an explicit learning structure; rather they compare the programmer's arrangement of an intended "logical" sequence with a "scrambled" sequence of program frames. This study (Experiment I and Experiment II) sought to determine, for practical instructional design purposes, how learning is affected by an instructional sequence of units derived from, and based on, the transfer implications underlying a task analysis and learning hierarchy.

Experiment II

Considering that the forward sequence group took significantly less time to complete the program and was superior on final task (posttest) performance, continued research related to this study is believed worthwhile. Also, the implications for low I.Q. Ss should be pursued. Also, since lower ability Ss benefitted more from the forward sequence than did high or average ability Ss, research should be conducted to see if younger Ss benefit more than do older Ss from a sequence based on a learning hierarchy.

Comparison of Results for the Two Experiments

Summative evaluation. It was expected that the forward version of the program would result in better performance for the older and more advantaged children in Experiment II (Exp. II) than for the younger and less advantaged Ss in Experiment I (Exp. I). This appears to be borne out for some dependent variables, but not for others.

In agreement with expectation, Ss in Exp. II did take less time than Ss in Exp. I, and Ss in both Experiments did take less time to learn in the forward order than in the random. Also, fewer program errors were made in Exp. II than in Exp. I. (See Table 2).

The competency test data cannot be directly compared for the two experiments, since two changes were made between Exp. I and Exp. II. In Exp. I the test frames were not labelled as a test, and feedback was provided. In Exp. II the test frames were labelled "Test Question," and feedback was not provided. This change was made in Exp. II to provide more "testlike" data for the planned path analysis, discussed later. Had this change not been made, it is possible that the means would be higher in Exp. II than in Exp. I, which is not the case in the present data.

On the posttest, the means are all in the expected order: (a) Forward Sequence, Exp. II; (b) Forward Sequence, Exp. I; (c) Random Sequence, Exp. II; (d) Random Sequence, Exp. I. On this measure, then, the summative evaluation of the program, as reflected in Table 2, does show higher performance in Exp. II than in Exp. I, and higher performance in forward than in random groups. Even so, the desired "design criterion" of performance was not reached, illustrating how laborious indeed is the task of achieving "learning to mastery," even when a program has undergone three revisions.

One of the most unexpected findings is reflected in the retention test data in Table 2. The rank order of means is: (a) Random Group, Exp. I; (b) Forward Group, Exp. I; (c) Random Group, Exp. II; (d) Forward Group, Exp. II. This suggests an inverse relationship (among groups) between posttest and retention test, while within groups, there is a positive correlation between posttest and retention test in Exp. II. It is possible that this reflects a complex interaction effect of time to learn, I.Q., and treatment. In any event, if retention were the sole criterion, it would be necessary to emphasize heavily that the highest retention test mean was for the Random Sequence in Exp. I. The combined retention means for both experiments are higher for the random sequence. Possibly involved in this almost uninterpretable finding are the following: posttest scores, in which low I.Q. Ss but not other Ss profit from the forward sequence; the tendency for low I.Q. Ss to take more time; the lower correlation between I.Q. and posttest score for the forward than for the random group; and the apparent discrepancy between correlations within groups of posttest and retention test (.58 forward; .74 random) and the inverse rank order of means among groups for the two measures.

In summary, the program, in its present form yields highest posttest scores for the older, more advantaged Ss in Exp. II who had the forward sequence, but it yields the highest retention scores for the younger, less priviledged Ss, who had the random sequence in Exp. I,

who also took the greatest learning time. So the condition that was best for immediate retention was not best for delayed retention.

Treatment effects. Rather poor agreement between the two experiments was found on treatment effects, in terms of statistically significant differences, although many trends are the same for the two sets of data. Of course, one reason for conducting Exp. II was that greater reliance would be placed on the data, due to the larger Ns and to the more appropriate Ss. On this ground, one would expect to find more evidence of treatment differences in Exp. II than in Exp. I, considering the results expected on basis of the theory underlying the hypotheses stated prior to Exp. I. This expectation was realized in the results.

In Exp. II, the forward treatment was superior to the random treatment in learning time (.01) and in posttest scores (.05), while in Exp. I there was a significant difference only on the program error rate, in favor of the forward treatment, when transformed scores were used. The conclusion is that forward is superior to random sequence in learning time and posttest score, but not in retention score. Inspection of the means, however, indicates that this superiority is largely due to the results for low I.Q. Ss. Overall, the magnitude of the treatment differences gives some pause as to the hierarchical theory, especially considering the retention data. Complex interactions are suspected, making practical application of the findings to bring about large learning improvements a very difficult task. Were it not for the results for the low I.Q. Ss, it would be difficult to defend the practicability of widespread application of the results. However even this caution needs qualification: since the methods used in development of the program led to classification of frames by competency, to insure that there was substantial, relevant material for each unit, this probably did result in a more adequate program than if this procedure had not been followed, and this, in turn, could lower the size of the differences among means for treatments. It may be recalled (from Volume I) that this sorting of frames by competencies for the Princeton Algebra Program led to the discovery that there were no frames (or few frames) for some competencies. Gagné and his associates called attention to the rather low performance resulting from use of that program. So it may be that the more carefully and thoroughly the program units are developed, the less the difference among sequence variations. Also, the units involved an intact sequence of frames, within unit, across treatments. This, of course, isolates unit sequence effects from frame sequence and relevant content, thus removing these two program characteristics as independent variables. So effects of unit sequence alone could be expected to be smaller than if both unit and frame sequence and content relevance were varied.

In conclusion, the total program development procedures employed gave a within-unit logical sequence of frames which was the same for all unit sequence treatments. Thus the program, even in the reverse

and random unit orders, may be more adequately sequenced (within units) and may contain more adequate material than may often be the case. With all these considerations in mind, the size of the treatment differences is put in somewhat different perspective. Even so, the fact that the design criterion was not reached, remains, leaving room for further program improvement. But this is stressed here neither to apologize for the data nor to avoid them--it is stressed to call attention to the need to continue doing such research with programs whose degree of effectiveness is known, and is recognized as an experimental variable. One could speculate at this point how the results might differ with programs of various degrees of effectiveness and various degrees of internal structure (frame sequence within unit). Said differently, the reverse and random sequences are not entirely unstructured, and frames were added and revised, seeking for adequate, relevant content. The retention test results suggest the need to add more practice frames into the program to achieve overlearning, particularly in view of the demanding nature of the posttest and the retention test (see Volume I for a discussion of this).

Interaction results. In Exp. I, there were no significant interactions. In Exp. II the only significant interaction was between sequence and I.Q. upon attitude score.

Consistent trends, but not all of them significant, were found, in both experiments, suggesting that low I.Q. Ss profited from the forward sequence on all dependent measures: time; program errors; competency test; attitude; posttest; and retention test. It would appear that middle and high I.Q. Ss are able to either learn directly from a random unit sequence, or are able to recall and restructure material as they study a random sequence. The time increase for such Ss may not be due to "difficulty" in reading and understanding, but to "time out to recall, restructure, and figure out" the correct sequence. An alternate hypothesis is that high I.Q. Ss do not need to mentally resequence the task in order to learn. By a greater generalizing and problem-solving capacity, they may do for themselves what the forward sequence treatment does for low I.Q. Ss, or, they may learn competencies simultaneously or in a changed order, but "put it all together" when facing the final task. Nevertheless, for the entire I.Q. range, the features of the forward program were effective for immediate, but not delayed retention. Only one of the six subgroups had a higher retention than posttest score for the forward group, while four such score increases are noted for the reverse group.

Conclusions

In Exp. I, the group having the forward, or hierarchically-based unit sequence, performed better only in terms of program error rates, than did groups having the units in reverse or random order. (Frames within units were in the supposedly logical order for all treatments).

In Exp. II, there were treatment effects on two dependent variables, and one interaction. The forward group took less time, making this the most efficient group, while also scoring highest in effectiveness on the posttest. An interaction of treatment and I.Q. upon attitude, was also found.

However the superiority of the forward sequence was not borne out in the retention test, given three weeks after the posttest.

For low-ability Ss, there were consistent trends, in favor of the forward group on all dependent variables. These Ss appeared to profit relatively more than did those of higher I.Q. from the forward sequence.

While the program did not meet the pre-set design criterion, the careful attention to content to match the competencies assumed needed for the task probably resulted in a program superior to less structured materials. It is believed that addition of more practice frames would improve both the learning and retention of the task, although it would also increase learning time, apparently a contributor to retention.

Complex interactions are suspected among I.Q., learning time, and program sequence, in considering how the difference between the results for the posttest and retention test may have arisen. More exploration of such interactions is needed to show how to obtain larger gains by sequencing according to a hierarchy. Apparently different sets of conditions contribute to short term vs. long term retention.

The fact that the sequence of frames was left intact, within units, across treatments, probably tended to minimize treatment differences, compared to a situation in which both frame and unit sequence would be varied.

Recommendations for Future Research

Investigation of triple interactions has been suggested as one avenue of further exploration, to gain a better understanding of how the results on retention could be improved.

It is suggested that for bright Ss, content relevance may be a more powerful variable than sequencing. Adequacy of the groups of frames could be varied to compare the effects of content vs. sequencing of content.

Since the forward sequence benefitted low I.Q. Ss the most, it would be desirable to further study sequence as a function of age of the learner.

Finally, it has been suggested that the empirical technique reported by Resnick and Wang (1969) could be employed to improve sequencing of units. It would be possible to take the data for the competency test from Exp. II in this study, and utilize results of path analysis techniques to infer a different sequence, which could then be compared in effectiveness with the logically derived sequence used in the present study. Since frames within unit are intact for all versions of the present program, only rearranging the pages would be needed to compare the two hierarchies in actual effectiveness.

Continued research is needed in how to construct effective instructional programs. Introducing variations in content from the program used in this study could be made, utilizing the present data and program as a baseline for determining which variations improve effectiveness. Protracted research, using variations of the same program, presents one attractive strategy for planning future research.

TABLE 1

F Ratios for Analyses of Variance: Sequence Variations, Ability Levels, and Sequence by Ability Interaction[1]

Source of Variation	df***	Time to Comple- tion	Program Errors	Competen- cy Test Score	Attitude Measure	Posttest Score	Retention Test Score
Sequence Variations	II 1	6.99**	2.51	.04	1.27	4.11*	.04
	I 2	.68	4.08*	.75	.16	.64	.29
Ability Levels	II 2	6.78*	8.93**	16.88**	2.22	18.54**	15.43**
	I 2	1.08	6.87**	3.30*	4.62*	3.08	3.77*
Sequence by Ability Interaction	II 2	.31	.95	1.22	3.42*	2.70	.47
	I 4	.33	.27	.30	.48	.92	1.29

* $p < .05$
** $p < .01$

*** In each row, "II" stands for Experiment II, and "I" stands for Experiment I.

[1] These data are based on raw scores for Exp. II, and on the square root transformation scores for Exp. I, as explained in Volume I of this report, due to heterogeneity of variance in Exp. I.

TABLE 2

Raw Score Means and Standard Deviations

Treatment Group*	Ability Level*	N	Exp.	Time to Completion Mean($\bar{x}$)	Time to Completion S.D.	Program Errors Mean($\bar{x}$)	Program Errors S.D.	Competency Test Score Mean($\bar{x}$)	Competency Test Score S.D.	Attitude Measure Mean($\bar{x}$)	Attitude Measure S.D.	Posttest Score Mean($\bar{x}$)	Posttest Score S.D.	Retention Test Score Mean($\bar{x}$)	Retention Test Score S.D.
Forward	High	32	II	119.37	26.36	9.00	6.56	8.12	1.31	74.22	11.58	6.75	3.40	5.80	3.90
N		7	I	177.86	51.95	15.00	19.03	9.14	.69	77.14	13.18	5.43	3.41	5.43	3.91
	Average	31	II	121.93	23.05	12.03	10.00	7.29	1.72	70.00	8.27	5.64	3.73	4.00	4.15
II-92		7	I	193.57	46.16	28.43	17.73	7.71	1.89	58.57	9.88	4.29	1.11	1.86	1.77
I-20	Low	29	II	144.83	48.74	15.41	9.49	6.45	2.05	73.45	15.65	4.18	4.16	2.46	3.44
		6	I	211.67	45.46	28.83	24.18	8.17	1.47	67.50	8.22	5.33	3.14	6.17	3.82
Random	High	30	II	132.83	37.45	11.53	8.16	7.83	1.34	77.83	12.71	6.67	3.41	6.61	3.71
N		6	I	217.50	54.75	23.17	4.53	8.50	1.52	70.83	8.61	6.33	2.16	6.83	2.48
	Average	25	II	142.00	35.21	11.80	7.78	7.80	1.44	77.60	12.34	5.32	4.22	4.17	4.58
II-83		7	I	203.57	39.97	40.86	12.80	7.71	1.11	57.86	10.35	3.71	1.60	4.29	3.30
I-20	Low	28	II	154.29	39.08	20.61	17.22	6.07	1.96	68.75	14.76	1.26	2.31	1.85	2.49
		7	I	225.71	102.45	67.00	40.32	7.00	2.31	70.71	18.58	2.57	3.10	3.29	3.35

* In each row, "II" stands for Experiment II, and "I" stands for Experiment I. The total N's for the two treatments are shown in the first column; the N's for ability groups within treatments are shown in the second column.

TABLE 3

Intercorrelations Among Variables: Forward Group

	I.Q. Score	Pre-test	Time to Comple-tion	Pro-gram Errors	Compe-tency Test	Atti-tude Measure	Post-test	Reten-tion Test
I.Q. Score		.16	-.28	-.35	.40	-.04	.32	.38
Pretest			-.14	-.18	.13	.06	.26	.36
Time to Completion				.15	.06	-.02	.06	-.15
Program Errors					-.25	-.11	-.36	-.43
Competency Test						.04	.51	.48
Attitude Measure							.04	.20
Posttest								.58
Retention Test								

TABLE 4

Intercorrelations Among Variables: Random Group

	I.Q. Score	Pre-test	Time to Comple-tion	Pro-gram Errors	Compe-tency Test	Atti-tude Measure	Post-test	Reten-tion Test
I.Q. Score		.13	-.23	-.27	.40	.28	.58	.55
Pretest			-.18	-.13	.18	.09	.33	.38
Time to Completion				.16	-.19	-.22	-.18	-.31
Program Errors					-.45	-.21	-.31	-.37
Competency Test						.15	.53	.47
Attitude Measure							.25	.19
Posttest								.74
Retention Test								

REFERENCES

Briggs, L. J. Sequencing of instruction in relation to hierarchies of competence. Pittsburgh, Pa: American Institutes for Research, 1968. Monograph No. 3.

Briggs, L. J. Handbook of procedures for the design of instruction. Pittsburgh, Pa: American Institutes for Research, 1970. Monograph No. 4.

Esbensen, T. Working with individualized instruction. The Duluth experience. Palo Alto, Cal: Fearon Publishers, 1968.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M., & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs: General and Applied, 1961, 75. (Whole No. 518).

Gagné, R. M., and Staff, University of Maryland Mathematics Project. Some factors in learning nonmetric geometry. Monographs of the Society for Research in Child Development, 1965, 30, 42-49.

Levin, G. R., & Baker, B. L. Item scrambling in a self-instructional program. Journal of Educational Psychology, 1965, 54, 138-143.

Mager, R. F. On sequencing of instructional content. Psychological Reports, 1961, 9, 405-413.

Niedermeyer, F. C., Brown, J., & Sulzen, B. Learning and varying sequences of ninth-grade mathematics materials. Journal of Experimental Education, 1969, 37, 61-66.

Payne, D. A., Krathwohl, D. R., & Gordon, J. The effect of sequence on programmed instruction. American Educational Research Journal, 1967, 4, 125-132.

Resnick, L. B., & Wang, M. C. Approaches to the validation of learning hierarchies. Paper presented at the Eighteenth Annual Western Regional Conference on Testing Problems, San Francisco, May, 1969.

Roe, K. V., Case, H. W., & Roe, A. A scrambled versus ordered sequence in autoinstructional programs. Journal of Educational Psychology, 1962, 53, 101-104.

Spencer, R. Y. Application of a learning hierarchy to sequence an instructional program, and comparison of this program with reverse and random sequences. Unpublished doctoral dissertation, Florida State University, Tallahassee, 1971.

Final Report

Project No. 1-D-007
Grant No. OEG-4-71-0071

Rosemary Y. Spencer
Department of Educational Research
College of Education
Florida State University
Tallahassee, Florida 32306

Application Of A Learning Hierarchy To Sequence An Instructional Program, And Comparison Of This Program With Reverse And Random Sequences

September, 1971

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development
(Regional Research Program)

Final Report

Project No. 1-D-007
Grant No. OEG-4-71-0071

Application of a Learning Hierarchy to Sequence An Instructional Program, and Comparison of This Program With Reverse and Random Sequences

Volume I of II Volumes

Rosemary Y. Spencer
The Florida State University
Tallahassee, Florida

September, 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

THE FLORIDA STATE UNIVERSITY

COLLEGE OF EDUCATION

APPLICATION OF A LEARNING HIERARCHY TO SEQUENCE AN INSTRUCTIONAL PROGRAM, AND COMPARISON OF THIS PROGRAM WITH REVERSE AND RANDOM SEQUENCES

By
ROSEMARY Y. SPENCER

A Dissertation submitted to
the Department of Educational Research
in partial fulfillment of the
requirements for the degree of
Doctor of Education

Approved:

____
Professor directing Dissertation

____

____

____

____

August, 1971

____
Dean, College of Education

APPLICATION OF A LEARNING HIERARCHY TO SEQUENCE AN INSTRUCTIONAL PROGRAM, AND COMPARISON OF THIS PROGRAM WITH REVERSE AND RANDOM SEQUENCES

(Publication No.)

Rosemary Y. Spencer, Ed.D.
The Florida State University, 1971

Major Professor: Leslie J. Briggs

The purpose of this study was to test the hypothesis that a learning program which is sequenced as implied by the task analysis of the instrucitonal objective, as shown in the "learning hierarchy" for the task, will result in superior learning as compared with programs sequenced in the reverse order or in a random sequence. To test this hypothesis, three stages of work were required.

First, it was necessary to find or develop an instructional program sequenced according to an explicitly-stated hierarchy. At the outset of the study, there was no suitable learning program available in which the instructional sequence was derived from the subordinate competencies within a hierarchy. Consequently a program was selected for redesign into an instructional sequence which would be based on the competencies of a learning hierarchy. The objective of this program represented a meaningful curriculum topic (task) in algebra for which a hierarchy was already available. This

program was revised and resequenced according to the hierarchy by classifying and grouping together the individual program frames representing the instruction for each competency in the hierarchy.

Second, in order that research to test the sequencing hypotheses would be based on an effective rather than an ineffective program, the experimental program was subjected to three tryouts and three revisions in an effort to establish teaching effectiveness at each competency level and for the final task to the 85/85 design criterion (85% of the pupils would score 85% or better on criterion tests). One tryout was made with individual students (Ss), one tryout with a small group of Ss, and one tryout with another larger group of Ss. Revisions were made between tryouts (formative evaluations), and the final tryout was taken as the summative evaluation, showing the extent to which the design objective was met before the program was used in the experiment.

The tryouts and revisions of the program succeeded in achieving: (a) a 35% reduction in the average time to complete the program; (b) lowered program frame errors by 42%; and (c) a 48% gain in the number of competencies mastered. There was, however, a decrease of .78 of a score from the small group posttest mean to the summative evaluation posttest mean. For the summative evaluation, Ss took an average of 3 hours and 20 minutes to complete the program. The program frame error rate was 15%. The summative evaluation resulted in a competency test standard of 85/79, and posttest results for the final task reached a standard of 67/49.

The third stage of the study was the conduct of the experiment comparing the hierarchy (forward) sequence, with reverse and random sequences. The sequence of frames within the instructional units of the program was kept constant, while sequence of presentation of instructional units was varied. It was hypothesized that the forward sequence group would be superior to the other two groups on the following dependent variables: (a) time to complete the program; (b) number of errors on program frames; (c) mastery of subordinates competencies; (d) attitude toward the program; (e) mastery of the task; and (f) retention of the task mastery. The effects of the sequence variations were studied for high, average, and low mental ability levels. Interaction was expected between sequence and mental ability level upon the dependent variables.

Analysis of the dependent variables on regular Biomedical 05V computer program resulted in nonsignificant results, and revealed the variability to be heterogeneous rather than homogeneous. A square root score transformation analyses of variance resulted in one significant difference, that of number of program frame errors, in favor of the forward group. There were no significant interaction effects found for sequence and mental ability upon performance. However, the low ability forward group took less time to complete the program, had fewer program errors, showed greater mastery of subordinate competencies, showed greater mastery of the task, and had better retention of the task mastery, than either the

low ability reverse or low ability random groups. Also, the low ability forward group had better retention of the task mastery than did the high ability forward group.

Further research may verify the implication from the above trends in results, namely that lower ability Ss receive greater instructional benefit from a careful instructional sequence than do high or average ability Ss who are better able to resequence and reorganize material for themselves. It is also likely that a second administration of the forward version of this much-revised program to a more advantaged sample of learners would reveal that the program did fully meet the original design objective.

ACKNOWLEDGMENTS

The writer is deeply grateful to her major professor, Dr. Leslie J. Briggs, who directed the study, and to the other members of her committee, Dr. Jacob G. Beard, Dr. Robert M. Morgan, Dr. Mildred E. Swearingen, with thanks to Dr. F.J. King for his assistance with the statistical analyses.

Acknowledgment is given to the Office of Education Regional Research Program for funding of later stages of effort on this study, and the Office of Education (Project No. 1-D-007, Grant No. OEG 4-71-0071) is also supporting a follow-up study described in the last chapter of this dissertation.

Appreciation for permission to conduct this research in the Leon County Public Schools is given to Dr. Ruth S. Mitchell and the other members of the Leon County Advisory Committee on Curriculum Research. Without the friendly cooperation of Mrs. Emma Wade, classroom teacher and Mr. Devurn H. Glenn, principal at R. Frank Nims Middle School, this study could not have been completed.

Finally, the writer thanks her husband, Platt, for his unfailing encouragement and support.

EXPLANATORY NOTE

The federal support of portions of work reported here began subsequent to the development of the instructional materials. For this reason, the Princeton Algebra Program, and the Basic Algebra Skills Learning Program mentioned on page viii in the Table of Contents are not found in the Appendix of this report as the Table of Contents would suggest. This accounts for the lack of continuity in numbering of pages in the Appendix to this report. These deleted materials are, however, contained in the writer's dissertation which is available to the public.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE PROBLEM

Sequencing of Instruction

The primary purpose of this study was to investigate the importance of the variable of sequencing of instruction at a particular level of specificity in the planning of instruction.

Since not everything that a student (S) is to eventually learn in school by the twelfth grade can be taught at one time, all educators have had to deal in some fashion with decisions about what to teach first and what should come next.

At a very general level in curriculum planning, educators plan, for example, which sets of mathematics objectives should be mastered each year. Thus within each year's learning, more specific tasks must be ordered for the week-to-week learning. Then the teacher or textbook writer keeps pressing this sequence planning down to a day-by-day level, or to the even finer level of sentences in a textbook or steps in the group teaching of a single hour's lesson. At the other end of the scale, elementary school comes before high school, and so forth.

Thus the practical necessity for dealing with sequence issues is an inescapable task for all teachers. Hence the first self-evident fact is that all of education requires somebody to decide upon sequencing at the various levels of specificity. At one level this may be done by curriculum planners; at another level by teachers or developers of instructional materials, and at some levels by the learner or by the writer of "frames" in a programmed instruction unit.

Opinions vary among educators as to whether the teacher or the learner should make such sequence decisions, and upon the importance of sequencing as a variable in the instruction. Theorists would differ on the rationale upon which such decisions should be based, and research findings are not in agreement as to how much it matters which decisions are made or who makes them.

In an effort to confine this study to one particular level of specificity in dealing with sequencing, it was decided not to go into as great detail (as studies in programmed instruction have gone) as the "frame" level in comparing alternate sequences to test hypotheses, nor on the other hand to deal in the gross terms of entire courses covering a semester or a year of study. Instead, the decision was to work at the task level of detail, when a task is taken to mean not more than a few hours of instruction; e.g., not over 20 clock hours of study or teaching time, and also not less than 3 hours of study or instruction. These limits were set to distinguish this "level of sequencing" from

sequencing of "frames" which take only a few seconds or a minute or so, and from "course units" which may cover six weeks of instruction.

The decision to deal with the task level was made not only to define the level of specificity and to restrict the scope of the study but also to permit comparison of results with those of Gagné and his colleagues (Gagné & Paradise, 1961; Gagné, 1962; Gagné, Mayor, Garstens & Paradise, 1962; Gagné & Staff, 1965) who were investigating transfer of learning within the subordinate part of a learning task.

Learning Hierarchies as a Basis for Sequencing

As mentioned earlier, one might have different theoretical or rational grounds for how he makes his sequencing decisions. Some might sequence instruction in the historical order of the development of knowledge in a subject; some might appeal to the logic of a subject-matter outline; some might teach practical skills in the order in which they are performed in the world of work; some might just furnish many learning resources and materials and leave it to the learner to find his own way through them; some might arrange tasks according to the degree of complexity or difficulty thought to be involved; some might use an inductive strategy in which elements of instruction appear in different order than under a deductive strategy.

It would appear reasonable that the same designer might follow one of the above sequencing rationales for teaching motor skills, a different one for intellectual skills, and still a different one for affective objectives.

In this study, as in those cited above by Gagné and others, it was desired to utilize a learning hierarchy as the basis for the sequencing of instruction for an intellectual-skill task.

Since Gagné's work relating to hierarchies is reviewed in Chapter II, it will suffice here to say that, in a global sense, he accomplished the following:

1. He developed hierarchies to systematically infer the subordinate competencies which should presumably be learned before the learner could be expected to perform the final task located at the top of the hierarchy.

2. He arranged the subordinate competencies in layers by starting with the final task and asking "what would an individual have to know how to do in order to achieve successful performance of this class of task, assuming he were given only instructions?" The answer to this question defines one or more subordinate tasks, or competencies, which support the given final task. Then the question is asked of each subordinate task so defined, "what would the individual have to know how to do in order to be able to achieve this competency, when given only instructions?" This analysis is repeated for each subordinate

competency, going down the hierarchy, until the entire hierarchy is defined (Gagné & Paradise, 1961).

3. Once such hierarchies were so derived, he tested students (Ss) over the different competencies, noting each competency as passed or failed, and then either used the pass-fail data to verify the assumed directions of transfer, or he conducted instruction on the failed competencies, and then retested for mastery of the competencies or the task itself.

4. By such methods he found empirical support for the need for the component competencies identified and for their arrangement into layers as shown in the hierarchy.

5. He then comments on the implications of his findings for the sequencing of instruction on the various competencies. In the studies reviewed to date, as reported in Chapter II, he has not experimentally manipulated total sequences of instruction in such a way as to test the assumption that because transfer is from a lower level to a higher level that the competencies of the lower level should be taught before the higher level. His studies involved an unspecified teaching sequence, not revealing <u>when</u> each competency was learned, but just verifying that if the competencies of one layer are not learned, the competencies of a higher layer will probably not be learned. However, Gagné writes:

> . . . the theory of learning set hierarchy has a number of implications for the programming of productive learning. Chief among these is the idea of designing the frames of a program in such a way that they: constitute

> an ordered sequence logically related to the hierarchy of learning sets [competencies] of the desired final task, provide for recallability of subordinate learning sets, and furnish the guidance to thinking which will enable the learner to integrate subordinate learning sets in the performance of new tasks (Gagné & Paradise, 1961, page 16).

Briggs (1970), attempting to incorporate Gagné's work in a recommended procedure for the design of instruction, has recommended that after a learning hierarchy has been drawn up, one might number the competencies to show the teaching sequence proposed, to take account of transfer within levels as well as among levels. Transfer refers to the recall of previous relevant learning and use of this learning during the learning of something new. Briggs reasons that by sequencing instruction in the order implied in the hierarchy, each subordinate competency would be taught when it is most needed, thus providing for recall of needed previous competencies and the instructional guidance for thinking, to enable the learner to master the new competencies and then the final task.

Thus Briggs has attempted to show an interim way to use the implications of Gagné's findings in the actual design of instruction until future research reveals more clearly the actual closeness of the relationship between hierarchies and effective sequencing.

It was the purpose of this study to help reveal the extent to which this literal translation of Gagné's data into practice may find empirical support. To do so,

it was decided that a rather radical experiment was needed, as suggested by Briggs (1968, p. 7), in which one group follows the sequence implied by the hierarchy, another group follows an "inverted" (reversed) sequence, while a third (control) group follows a random sequence in study of the competencies of a task.

Hypothesis To Be Tested

The hypothesis tested in this study was that a learning program sequenced in accordance with the hierarchy (forward sequencing) will be more effective than a program which inverts this sequence (reverse sequence) or a program in which the competencies are taught in a random sequence.

It was hypothesized further that the forward sequence group would be superior to the other two groups on the following dependent variables:

1. time to complete the program
2. number of errors on program frames
3. mastery of subordinate competencies
4. attitude toward the program
5. mastery of the task
6. retention of the task mastery

The effects of the sequence variations were studied for high, average, and low mental ability levels. Interaction was expected between sequence and mental ability level upon the six dependent variables.

Experimental Approach

To test the above hypotheses, it was necessary to develop or to locate a hierarchy and a learning program for a specific task of the desired level of magnitude (taking between 3 hours and 20 hours of study time). It was then necessary to ascertain that the program was an effective one for teaching the task in its forward sequence. Then it was necessary to rearrange the instructional sequence for the competencies to institute the reverse and the random sequences.

Since a hierarchy and a program were available for a task of the desired magnitude, it was decided to start with them, and to revise them through formative evaluation procedures until the program met the desired level of effectiveness. Then the two alternate programs (reverse and random sequence) would be compiled simply by rearranging the sequence of blocks of intact program frames corresponding to the various competencies. Then the three programs could be compared in effectiveness, using the dependent measures defined above as criteria of effectiveness.

In summary, the following major stages of work were planned:

1. Sequence the instructional units of the program to correspond to the order implied by the competencies of the hierarchy.

2. Revise the program until it met a reasonable level of teaching effectiveness, as shown by empirical tryouts.

3. Develop measures for all specified dependent variables.

4. Conduct a summative evaluation of program effectiveness (forward version).

5. Prepare the two alternate program versions (reverse and random sequence).

6. Conduct the 3-group experiment to test the hypotheses.

The Experimental Task

In the search for an experimental task suitable for the purposes of the study, it was desired to select one actually a part of a school curriculum, thus making it possible to justify use of class time in a school for this research. It was also desired to keep to the lower end of the range in study time (from 3 to 20 hours).

These study time limits for the desired task were chosen to avoid a program so short that the content of the program frames could easily be remembered and mentally rearranged by the learner, and to avoid a program so long as to make the process of formative evaluation and program revision too time consuming. Also, this time range is practical to identify the amount of learning represented by a "task" as larger than a single program frame or competency, but smaller than a major "course unit."

Gagné and Paradise (1961) had previously worked with a task somewhat larger than the one sought for this study, but whose structure had been analyzed into a hierarchy with three main branches, at the top of each of which was a "subtask." Since the hierarchy indicated the three branches as independent ones, it was feasible to select one of the three as the task for this study.

In the remainder of this report, the designation "task I,1" refers to the portion of the larger task of Gagné and Paradise. It is this portion which is employed for the present study.

The entire task by Gagné and Paradise was "Solving linear equations," and the task I,1 is "Simplifying an equation by adding and subtracting terms to both sides." This latter skill is the "task" for the present study.

The subordinate competencies of the task are arranged in layers designated IV for the lowest layer, then going up the hierarchy, as layers III and II. For competencies on the same level (or layer), Arabic numbers are added, so that each competency has an identification, such as IV,1; IV,2; etc. A description of the competencies of task I,1, listed in the order in which they were taught in the forward group of the present study, is found in Appendis B.

CHAPTER II

REVIEW OF THE LITERATURE

Alternate Ways of Sequencing Instruction

Educators and researchers have differing views as to why sequencing of instruction is important, who should prepare the sequence, and how to provide effective instructional sequences. While almost all agree that sequencing of school learning is important, differences arise because of differing ways of conceptualizing learning.

Mager (1961) implies that the S can best select the order in which he should learn various skills, when reporting a study in which learner sequencing was different from that of course outline or textbook sequences.

Six persons with no training in electronics, but who expressed a desire to learn something about the topic served as Ss in Mager's experiment. Each S met individually with the instructor for several sessions. In each session, the learner was told that the content covered was dependent upon the questions he asked. There was no lecturing or reading. Transcripts of the sessions showed that: (a) each S entered the instructional situation with a rather large body of information about electronics, as revealed by their questions; (b) Ss wanted to start with simple wholes and progress to more

complex wholes; and (c) there was a commonality among the independent learner-generated sequences.

In a follow-up study (Mager & Clark, 1963), adult trainees were given a list of instructional objectives and a variety of learning materials, and allowed to follow their own time schedule in achieving the objectives in the order in which they chose. Learning time was reduced an average of 65% over the formal course, but there was no control group to show how the formal course might have been improved. Both of these experiments have pertinent findings for adult education, but would have to be repeated with additional features added to have relevance for school learning.

Skinner's views (1958) regarding the sequencing of programmed instruction seem to emphasize the importance of sequencing more than many other researchers. Actually, the requirement of small teaching steps is based on Skinner's theory of reinforcement to shape behavior. He argues for a sequence of discrete frame steps, by means of which the learner must respond correctly. Then confirmation of the correct response serves as reinforcement to shape behavior. Skinner does not emphasize transfer, the recall of previous learning to learn something new. Also, he is interested in the concept of shaping behavior rather than establishing conditions for various kinds of learning.

Pressey developed a method of instruction called adjunct autoinstruction for use with instructional materials, which helps the learner to structure the sequence of the learning materials for himself (1926).

By Pressey's method, the S is encouraged to "range over" headings and overall organization of the material before reading them. Then after he has read the material, for example a chapter in a textbook, he answers a series of questions about the reading in a "practice test" or "self-check test." The S can find out if he is right or wrong after each response made, thus his correct learning is confirmed, and he is shown which areas of the material require further study. Pressey's method is a combination of experimenter-determined control and learner-controlled sequencing.

Esbensen (1968) infers that an analysis of course content will reveal that one sequence may be as good as another. In preparation for individualizing the instruction of a Duluth, Minnesota public elementary school, the staff of teachers wrote objectives for the subject matter areas of mathematics, science, social studies and language arts. Each objective was categorized according to Bloom's (1956) taxonomy of categories: knowledge, comprehension, application, analysis, synthesis, and evaluation. The objective numbers were placed on the curriculum map from left to right suggesting a desirable sequence, with lines connecting any relationships of dependency. Due to the scarcity of lines connecting the objectives indicating a decided lack of dependent relationships among the objectives, Esbensen concludes that sequencing of subject matter may not be as important as was once thought. This lack of dependent relationships among the objectives points to the difficulty

of attempting to sequence subject matter content, rather than a sequence of skill competencies supporting a defined behavioral final task.

Experiments by Gagné and His Associates

The experiments reported in this section began with an explicit rationale of how productive learning takes place. Gagné and others present a learning hierarchy of learning sets (competencies) supporting a final task, and they state the direction in which transfer is predicted to occur among the subordinate competencies.

In an early study, Gagné (1962) derived a hierarchy for the task "finding formulas for sum of n terms in a number series," by beginning with this task and asking, "what would the individual have to be able to do in order that he can attain successful performance on this task, provided he is given only instructions?" This question was then repeated for each subordinate task so defined, to derive nine subordinate competencies for the final task. At the very botton of the hierarchy may be found the learner's relevant entering competencies brought to this learning from prior learning.

Beginning with the final task, Gagné administered test items over each competency to ninth grade boys. If a boy failed the final task, the next highest test item was given and so on, going down the hierarchy. When successful performance was reached for any given individual, a learning program for the next highest level previously failed was

administered. Then test items were given over the remaining lower levels. The results showed there were no instances when an individual was able to perform a higher-level competency if he was unable to perform a lower-level competency related to it. After completing the instruction for previously failed competencies, each boy was again tested, so that scores on tests given both before and after administration of the learning program could be compared. Comparison of these scores supported the idea of the learning hierarchy.

Gagné and Paradise (1961) corroborated their prediction of positive transfer from recalled competencies to higher competencies of a learning hierarchy, with proportions ranging from .91 to 1.00.

They first analyzed an existing program, the Princeton Algebra Program, to define a hierarchy of 22 competencies for the final task, solving linear algebraic equations. The procedure employed in the study was as follows: first, the learning program was administered to seventh grade mathematics Ss. After Ss had completed the program, a performance test was administered, followed immediately by a transfer test. The next day, a test covering all the competencies of the hierarchy was administered. The predicted positive transfer from lower to higher competencies was affirmed by noting the pass and fail patterns between competency test items.

The overall achievement on the performance test and transfer tests indicated that the learning program was only of low-moderate effectiveness.

Further evidence in support of the learning hierarchy theory was provided in a study by Gagné, et al. (1962). For this study, a learning program on addition of integers was analyzed, and a learning hierarchy was derived (according to the procedure described earlier) containing 12 subordinate competencies. Four parallel forms of the program were prepared, with variation in the amount of guidance, and the amount of repetition. After administration of the learning program to the 137 seventh graders, tests over task performance, transfer, and subordinate competencies were administered.

Acquisition of competencies at successively higher stages of the hierarchy was found to be dependent upon prior mastery of subordinate learning sets, with instances of confirmation ranging from 97% to 100%. The effects of the two programming variables, guidance and repetition, were small in comparison to the consequences of addition or omission of subordinate competencies.

Still another study confirmed Gagné's notion that the attainment of higher competencies in the hierarchy is dependent upon mastery of lower competencies (Gagné & Staff, 1965). A self-instructional program was developed in non-metric geometry for the task "specifying sets, intersections of sets, and separations of sets, using points, lines, and curves." The programming variables (a) variety of examples, and (b) passage of time between stages of learning were studied by preparing five variations of the learning program. Following the learning, a test of achievement was administered

to the 90 sixth grade Ss. While the importance of sequencing was again confirmed, the two programming variables had no evident effect upon the learning effectiveness of the program.

Hackett (1968) developed a hierarchically ordered language comprehension skills test, to determine if language skills in school children would exhibit an ordering compatible with the hypothesis of transfer from lower-levels to higher-levels. To do this, she first identified eleven language comprehension skills from a review of the literature. These were stated as behavioral objectives, and arranged in a hierarchy of listening and reading comprehension. Next, two parallel forms of a test were constructed for the hierarchy of language skills test items, and administered to 1,186 Ss in second, fifth, eighth, and eleventh grades.

The results of a chi square analyses were statistically significant in favor of the hierarchical nature of skills at all four grade levels studied. Also, evidence of transfer among the language comprehension skills was indicated by patterns of relationships from lower-level to higher-level skills at all four grade levels.

Coleman (1969) provided evidence in support of Gagné's hypothesis that mastery of a final task is dependent upon mastery of all the subordinate competencies of that task.

The task of comparing in the analytic mode of social studies was analyzed and found to be hierarchically structured. Then the instruction necessary to teach each competency was defined. Sixth grade girls received diagnostic testing to

determine the point at which each S was ready to begin instruction. The experimental group received instruction on those subordinate competencies they had failed, but not on the final task. The control group did not receive instruction for their deficient competencies. The result was a significant difference in favor of the experimental group's ability to perform the final task.

Thus, research generally substantiates Gagné's notion of a learning hierarchy. The studies of Gagné and others have confirmed that when the competencies are defined for a hierarchical task: (a) there is positive transfer from lower to higher competencies in the hierarchy; and (b) Ss who have not mastered one or more competencies cannot perform the final task, while those who have been instructed in all the competencies can perform the final task.

Scrambled Versus Logical Sequences

Studies comparing logical with scrambled instructional sequences generally report negative, or no-difference findings. Primarily, these studies have manipulated the sequence of program frames, rather than of instructional units.

In a discussion of the methodological weaknesses of studies comparing sequences of instruction, Niedermeyer (1968), points out the difficulty in applying a learning structure to a program frame sequence. He distinguishes clearly between a learning hierarchy of subordinate skills and a series of programmed instruction frames based on subject matter content.

Roe, Case and Roe (1962) administered a 71-item elementary probability program in logical and scrambled sequences to 189 college freshmen. There was no significant differences between the two program frame sequences, and the investigators concluded that the effect of sequencing would be a function of the length and complexity of the program, the nature of the task, and the ability level of the Ss.

Levin and Baker (1965) scrambled frames for a portion of a geometry program, and presented it in logical and scrambled frame sequences. They found no differences between the scrambled and the original sequence on performance, retention or transfer.

Wodtke, Brown, Sands and Fredericks (1968) reported two experiments where computer terminals were used and a comparison of random and logical sequences of instruction in two content areas were made.

In the first experiment, a 74-frame program on number bases was administered to 80 education majors at Pennsylvania State University. This program was felt to contain a conceptual hierarchy where sequence would be important. There was a significant difference in program error rates in favor of the logical sequence, but there were no achievement differences or aptitude-sequence interactions with respect to posttest performance.

The second experiment utilized a program teaching discrete facts relative to the anatomy of the ear. No effect for sequence was hypothesized for the material was

apparently non-hierarchical or flat in structure. There were no differences in error rate or for posttest performance between sequence groups in this second experiment.

One of the better studies comparing a scrambled with a logical frame sequence, was reported by Payne, Krathwohl, and Gordon (1967). They hypothesized that the affect of scrambling would be greatest for those programs dealing with topics having the most internal logical development. The internal logical development of three programs in educational measurement was judged. These programs were administered to college sophomores in both logical and scrambled frame sequences. The results were, that all the scrambled forms of the programs had higher error rates, but other expected effects on performance and retention from scrambling did not materialize. They suggest that Ss of a sufficient age, particularly for a program of only brief duration, may be able to restructure the sequence of the material for themselves.

Use of the Hierarchy to Sequence Instruction

In a comprehensive review of rationales and experimental procedures in sequencing of instruction, Briggs (1968) states the need for experiments which distinguish between sequencing of frames and sequencing of competencies, or units of instruction. He further cites the need for evidence as to how sequencing of instruction is affected by a task analysis and learning structure. In this report, Briggs gives criteria for effective study of instructional sequencing.

These are: (a) use of a task of intermediate size; (b) a learning structure defined following Gagné's procedure for arriving at a hierarchy; (c) an instructional sequence developed from the learning structure; and the effectiveness of the instructional materials established for (d) each competency level, and (e) the total task.

An experiment which attempted to sequence an instructional program according to a learning structure, was reported by Niedermeyer, Brown and Sulzen (1969). They administered the Number Series Program from Gagné and Brown's 1961 study to 9th grade algebra Ss in logical, scranbled and reverse sequence versions. The logical group made significantly fewer program errors and performed significantly better on a test of concepts and a problem solving test; but there was no difference on a performance posttest. They concluded that for short programs sequencing may not be as crucial to cognitive outcomes as has been thought, since these Ss integrated and organized information regardless of the sequence of presentation. Unfortunately, it later became apparent that the Number Series Program was developed prior to Gagné's derivation of a learning hierarchy.

In contrast with other studies of instructional sequence based on Gagné's theory of the structure of learning, this study: (a) made use of a learning hierarchy to develop the instructional sequence; and (b) studied the affect of sequencing by varying program presentation of instructional units, rather than program frames.

CHAPTER III

METHODS

Overview of the Procedures

In order to test the hypotheses, presented in Chapter I, concerning the implications of a learning hierarchy for the sequencing of instruction, three distinct stages of work were required. These three stages, and the rationale for them, are summarized in this overview of procedures employed. The remainder of this Chapter supplies additional detail concerning the procedures.

First, it was necessary to find or develop an instructional program sequenced according to an explicitly-stated hierarchy, so that the sequence of the units of instruction could then be manipulated to provide the forward, reverse, and random sequences needed to test the hypotheses. It was decided that the format of programmed instruction provides a convenient vehicle for this research, since (a) individual program frames can be classified to match specific competencies in the hierarchy, and (b) the sequence of frames within units can easily be kept constant among the three versions of the program while varying the sequence of presentation of the intact units (groups of frames). Thus the teaching sequence

<u>within units</u> is the same for all three experimental groups, while the order of presentation <u>among units</u> can be varied to constitute the forward, reverse, and random treatments required for the experiment. Since it was desired that the program chosen represent a meaningful curriculum objective, thus justifying use of class time in a school for the experiment, several existing programs and hierarchies previously used for laboratory investigations had to be ruled out. As will be seen in the next section of this Chapter, the result was that the program finally used in the experiment was a drastically revised program on a topic (task) in algebra, for which a hierarchy was already available.

Second, it was desired that the experimental program be an <u>effective</u> one, so that research to test the sequencing hypotheses would be based on an effective rather than an ineffective teaching program. Since a frame-by-frame analysis of the original program indicated that it was not adequate for teaching all the assumed supporting competencies for the task, as reflected in the hierarchy, extensive revisions were made to insure that there were sufficient instructional frames for each of the competencies, and tryouts of the program were conducted in order to improve the effectiveness of the instruction. It was the objective of these tryouts and revisions to bring the program up to the design criterion of 85/85 (85% of pupils would score 85% or better on a test over the final task). But since the skill of the

programmer and the ability of the tryout Ss are always variables influencing the achievement of any stated design criterion, it is a matter of economy to decide when to accept a somewhat lower level of achievement rather than expending disproportionate amounts of time, effort, and money to close the gap between obtained and desired program effectiveness. For purposes of the present study, therefore, it was decided to accept whatever program effectiveness could be reached by three revisions and evaluations of the program. It was therefore decided to conduct one tryout with individual Ss, one tryout with a small group of Ss, and one tryout with a larger group. Revisions were to be made between tryouts (formative evaluation), and the final tryout was to be taken as the summative evaluation, showing the extent to which the design objective was met before the program was used in the experiment involving the three sequences of program units (forward, reverse, and random). Thus between the first step of choosing a program and a task hierarchy, and the final step of conducting the three-group experiment, these intermediate steps had to be taken:

1. Frames in the existing program were classified, using the description of the competency and the test item for each of the 10 subordinate competencies of task I,1 of the hierarchy (Gagné & Paradise, 1961), thus forming instructional units for the competency levels of the hierarchy. Each competency test item was placed at the end of the teaching frame sequence for that competency.

2. Identification of those competencies for which new frames were to be written, or old frames revised.

3. Construction of a 10-item test for the final task to serve respectively as pretest and posttest (see Appendix C).

4. Pretest administration to the five classes to take part in the study.

5. Tryout of the revised program with individual Ss drawn from high, average, and low I.Q. ranges.

6. Making second revision of the program, using student comments, frame responses, program errors, and competency test score from the first (individual) tryout.

7. Administration of the second revision of the program to three Ss in a small group, and administration of the posttest, followed by an individual interview with each S.

8. Final revision of the program, using responses of Ss to frames, information gained from the interviews, competency test items, and posttest items.

9. Administration of the final (third) revision of the program to another group as the summative evaluation.

10. Construction of a parallel form of the test over the final task to serve as the delayed retention test (see Appendix E).

11. Construction of an attitude questionnaire to measure S attitude toward sequence presentation for the experiment (see Appendix F).

Third, the final revision of the program was administered as the forward version of the program; and simply by rearranging program pages, the reverse and random versions were assembled for administration to the other two groups. Results of the formative and summative evaluation and of the experiment are presented in the Results Chapter.

The remainder of this Chapter provides added detail concerning these three stages of work.

Subjects

In the Frank-Nims School in Tallahassee, Florida, a teacher of five eighth-grade classes agreed to participate in the study. One class, her "homebase" class, was designated as the group from which Ss for three tryouts of the program would be drawn, by procedures described later in this Chapter. The remaining four groups were later each divided into three subgroups for assignment to the three experimental treatments. In these four classes, individual Ss were randomly assigned to treatments, stratified by I.Q. level as required by the experimental design, thus avoiding the sampling problems incident to assigning intact classroom groups to treatments.

This school may be characterized as having a large proportion of disadvantaged pupils; 39% of the children in the five classrooms scored less than 80 in I.Q. While such a school may be considered not untypical of many urban schools, it certainly is not typical of the I.Q.'s found in schools outside disadvantaged areas.

Since the program chosen for the study dealt with a task in algebra, it was decided (before the experiment was conducted) to use data only for Ss scoring I.Q. 80 or above, on the ground that those below 80 typically do not take algebra in the ninth grade.

While the above decision was deemed appropriate for the experiment, the question remains whether the fact of such a large number of children below I.Q. 80 might create conditions leading to less academic success for those above 80 than might otherwise be expected. This factor must be borne in mind when considering the data from the summative evaluation (the measure of program effectiveness) and when considering the experimental results which might be expected in other S groups having higher proportions of I.Q. above 80. (The experimental results for the forward sequencing group in the experiment might be viewed, of course, as a better summative evaluation, since more Ss were involved, and the program was administered to entire, intact classrooms, rather than just the remaining Ss in the "homebase" classroom who had not participated in earlier tryouts of the program.)

Materials

This section first reports the procedures used in development of an instructional program sequenced according to a learning hierarchy. Recorded next are the steps taken

to prepare the program for use in an experiment through testing and revisions (formative evaluation). Finally, a description is given of the tests constructed, and the rationale for the attitude measure is set forth.

Development of the Instructional Program

At the outset of the study, there was no suitable learning program available in which the instructional sequence was derived from the usbordinate competencies of a hierarchy. However, the Princeton Algebra Program had been used by Gagné and Paradise (1961) to develop a hierarchy representing the learning of the task, Solving Equations. But since the instructional sequence of the Princeton Algebra Program had not been determined by the hierarchy developed later by Gagné and Paradise, the plan was to group the frames from that program to form an instructional unit for each competency level for task I,1 of the hierarchy.[1]

Gagné and Paradise gave a description of, and developed a test item for, each subordinate competency of the hierarchy. Both these test items and the competency descriptions stated in the hierarchy by Gagné and Paradise were used in the present study to identify the instruction necessary to teach each competency. The competency descriptions and representative test items, presented in the order

[1]The Princeton Algebra Program can be found in Appendix A.

in which the competencies were taught, can be found in Appendix B.

Initial development. The Princeton Algebra Program was studied in detail, and the program frames were sorted into groups to form instructional units based on the hierarchical competencies. When this sorting was completed, it became clear that the implications of the sorting were in agreement with the finding by Gagné and Paradise that the Princeton Algebra Program was of only moderately-low effectiveness. The sort revealed that for several competencies there were either no teaching frames or so few frames as to suggest that the program would be inadequate. This can be seen from Figure 1, a list of frames compiled from the sorting of frames from the original program.

The Princeton Algebra Program contains 232 frames. A total of 69 frames, identified by the sort as belonging to task I,1, in the hierarchy, were retained for use in the revised program for this study. Thus this revision made use of only 30% of the original program frames. In the case of competencies 9 and 10, program frames which did not provide for further teaching toward the competency, but would only serve as additional practice, were not used in the revision. The frames, thus omitted from Figure 1, for competency 9 are: 73-76, 78, 79, 81-86, 88, 91, 93, 95-97, 99 and 118. For competency 10, the omitted frame numbers are: 150, 151, 155, 163, 170, 172, 173 and 185.

Fig. 1--Frames compiled from the Princeton Algebra Program

Competency	Hierarchy Box	Original Program Frame Numbers	Number of Frames
1	IVA,1	1-14, half of 18, 25-28, 32, 33, 171	22
2	IVA,2	66-71	6
3	IV,2	146, 147	2
4	IV,3	none	0
5	IIIA,1	36	1
6	III,2	57-63	7
7	III,3	none	0
8	II,1	none	0
9	II,2	72, 77, 80, 87, 89 90, 92, 94, 98, 116, (frames 87 and 98 were combined)	10
10	I,1	148, 149, 152-154, 156, 162, 164-169, 174, 180-184, 186, 190	21
			69

Thus, the first program revision required extensive writing of new program frames to teach those competencies not taught, or inadequately taught, in the original program. A total of 136 new program frames were written, and added to the original 69, resulting in a 205-frame program sequenced according to the competencies of task I,1 of the hierarchy. This does not include the competency test items, which consisted of the last program frame of each instructional unit.

Formative evaluations. The program revised as described above to teach only task I,1 in the hierarchy, was subjected to three evaluative testings and three revisions. The first revision was administered to three individual pupils; after necessary revisions it was administered to a small group of three Ss; after final revisions it was administered to another larger group of nine Ss, and this final evaluation constituted the summative evaluation.

Three Ss were chosen from the "homebase" classroom for individual tryouts in the following manner. The I.Q. range of 80 and above, or 80-121 in this classroom was divided by three, thus forming three mental ability levels whose I.Q. scores were: high, 108-121; average, 94-107; and low, 80-93. Then the name of one S was randomly selected from each of three boxes containing the names of Ss in each of the three mental ability levels.

Three other Ss were chosen from the "homebase" classroom for the small group tryout of the program in the same manner as above. After these two drawings of names, nine Ss of I.Q. 80 and above remained in this classroom, and they

were used for the summative evaluation. This section will summarize the procedures used in each test of the program and will report the program revisions which followed the testing.

For the first individual tryout, then, one S was randomly selected (from the homebase classroom) from each of high, average, and low mental ability levels, to work through the program individually with the experimenter (E). It was explained to these three Ss that by working through the program aloud, many program inadequacies would be detected. Each S was encouraged to point out areas that were confusing, frames in which he was not sure of his response, portions not consistent with a concept in an earlier part of the program, and places where the program was too easy or seemed to "talk down" to him. Each program frame was presented on a "4 by 6" card, with the correct answer on the back. The S read the frame, responded aloud, and then looked at the answer on the back of the card. The E tabulated the responses, and made note of the time each S began and finished the program.

Figure 2 presents a summary listing of program revisions made at each competency level based on frame responses, S comments, program errors, and competency test items following the individual tryouts with these three Ss.

This second revision of the learning program included only 33 frames of the original 69 from the Princeton Algebra

Figure 2.--Program revisions following individual tryouts

Competency	Program Frames			Revision
	Deleted	Revised	Added	
1	24	0	0	reduced from 38 to 14 frames
2	8	0	0	reduced from 12 to 4 frames
3	9	0	0	reduced from 24 to 15 frames
4	1	0	6	increased from 34 to 39 frames
5	10	0	0	reduced from 18 to 8 frames
6	8	0	22	completely rewritten; increased from 20 to 34 frames
7	1	4	10	revised and added; increased from 10 to 19 frames
8	0	8	11	revised and added; increased from 8 to 19 frames
9	9	1	18	completely rewritten; increased from 14 to 23 frames
10	3	18	11	completely rewritten; increased from 27 to 35 frames
	73	31	78	

Program; or 15% of the new program. The total number of frames changed from 205 to 210.

In addition to the above revisions designed to teach the identified competencies more effectively, the decision had to be made after the individual tryouts of the program as to whether the existing (previously identified) hierarchical competencies needed to be more broadly defined to reflect required additional instruction, or whether new competencies needed to be added to the hierarchy. The course chosen was to assume that the existing competencies 6, 10, and 7 required additional instruction. The assumption of a requirement for additional instruction was not needed for the other competencies of task I,1 of the hierarchy.

The evidence of need for additional instruction for these three competencies was that, even though Ss mastered the lower competencies subordinate to 6, 10, and 7, they could not successfully complete the test items for competencies 6, 10, and 7. A diagnosis of the difficulty revealed that although the instruction provided for the lower competencies was sufficient at those levels, it was not inclusive enough to prepare for the teaching of competencies 6, 10, and 7. Appendix D gives a detailed explanation for the need for additional instruction and a description of the added instruction for competencies 6, 10, and 7.

This twice-revised program was next administered to three Ss in a small group situation. As before, Ss were randomly chosen from the same classroom from which Ss for the

previous individual tryouts were chosen, representing the high, average, and low mental ability levels. From this point on, the program was presented on standard 8 1/2" by 11" paper, rather than on the "4 by 6" cards of the previous (individual tryouts) administration. The Ss were made aware that they were helping to revise the program. They were asked to mark the areas in the program that were hard to understand, and they were told that the E would discuss these areas with them later. Ss were instructed in the mechanics of taking the program, but after they had started the program, no further help or clarification was given. Indications of a lack of motivation were observed the first day of the program administration, probably because Ss knew their performance would not affect their grade, and because unlike Ss involved in the individual tryouts, the program was not being administered directly by the E. This lack of motivation was apparently corrected the day following, when Ss were told that their teacher would see the results of the test they were to take following completion of the program. The starting and finishing time of each S was noted. After completing the program, Ss received a posttest.

Then the third program revisions were made based on (a) observation, (b) individual interviews, (c) program errors, and (d) competency test and posttest data. These revisions were as follows:

1. Frames missed by all three Ss were revised.
2. The format, not the content, was changed for competency test item 3.

3. The instructional content and frame format for competencies 6, 7, and 8, was clarified.

4. Competency 9 frames teaching transposing of arithmetic numbers, and those toward the end of the unit, were rewritten.

5. Competency 10 teaching on transposing terms, was rewritten.

In addition to the above revisions made prior to the final evaluation of the program (summative evaluation), competency test items which heretofore had simply been the last frame in each instructional unit, were identified as a "Test Question." Also, directions for taking the learning program in large groups, with less close observation and monitoring were written; and a motivational statement was prepared.

Summative evaluation.The program underwent a third and final evaluation with all those Ss in the classroom used for formative revisions of the learning program, who had not already participated in an earlier evaluation of the program. Of the remaining 24 Ss, 6 were not used due to illness, suspension or unavailability of I.Q. score; thus a total of 18 Ss took the final revision of the program.

The cooperating teacher requested the assistance of the E in administering the program. The program directions were read aloud, and Ss evidenced no problem understanding the instructions. A "motivational statement," which was also read aloud, explained that the Ss would receive 100 points to buy classroom favors, and that their teacher would receive their test results. The Ss were monitored

by both the classroom teacher and E. As the Ss finished their programs, the time of completion was noted, and they received a posttest.

The planned time schedule for development of the program sequence, and formative revisions and summative evaluation of the program, was for a period beginning mid-September, to finish the first week in December. However, due to extensive revisions of the program, the summative evaluation was conducted the week prior to the Christmas holiday. A surprise assembly called the second day required that slower Ss, and those who had been absent, complete the program after Christmas vacation, thus lengthening intervals for both learning and testing.

Since the purpose of the summative evaluation was to establish the teaching effectiveness of the program for pupils likely to take algebra, prior to using it to test the experimental hypothesis, the decision was made to use data only for Ss scoring I.Q. 80 or above. Due to the restriction to I.Q. 80 or above, the number of Ss whose data were used in the summative evaluation was only nine (of the remaining 18 Ss). The results of these nine Ss who participated in the summative evaluation of the learning program are presented in the Results Chapter.

Description of Tests

Four tests were used in the study, a competency test of the subordinate competencies of the hierarchy, pretest and identical posttest for the final task I,1, and a parallel form of the pretest and posttest administered as a test of retention.

The competency test and pretest and posttest were used during the formative revisions and summative evaluation of the program. All four tests were administered for the experiment: competency test, pretest and posttest, and the retention test. An explanation of the purpose of each test, and how each was constructed, is now presented.

Competency test. The competency test consisted of the representative test items used by Gagné and Paradise (1961) in defining the competencies of task I,1 of the solving linear equations hierarchy. There were 10 items, one for each competency level (see Appendix B). The test for the subordinate competencies served a three-fold purpose in the study. First, it was used along with the competency description, during the initial development of the program, to classify program frames to form instructional units to coincide with the hierarchical competencies. Second, during the formative revisions and summative evaluation of the program, the competency test evaluated the effectiveness of the program at each competency level. Third, the competency test score was a dependent variable in the experiment, on the basis of the hypothesis that Ss who received the forward sequence, --that is, had mastered the lower competencies prior to introduction to the higher competencies,--would perform better on successive competency test items, than those Ss who were taught these competencies in a reverse or random order. The competency test was included within the program as the last frame of each instructional unit.

Pretest and posttest.The pretest and identical posttest were used to evaluate performance on final task I,1. The pretest administration was needed to ensure that Ss had not already mastered the final task. During the formative revisions and summative evaluation of the program, the posttest indicated program effectiveness in teaching for the final task. For the purposes of the experiment, posttest results were used for comparison among the forward, reverse, and random groups. The 10-item pretest and posttest was constructed by writing test items, varying in complexity, but all similar to the representative item used by Gagné and Paradise (1961) to define task I,1.

The posttest (see Appendix C) was a demanding test for several reasons. The simplest of the 10 problems required the student to perform at least five steps to arrive at a solution. The number of steps needed to solve a problem made simple arithmetic errors more likely. Toward the end of the test the complexity of the items was such that if Ss were careless in arriving at any of these steps the item would be incorrect. Consideration of the content of an algebra test, with principle-type level of learning, in itself indicates the demanding nature of the posttest.

Retention test.The retention test (see Appendix F) consisted of 10 items which were considered to represent a parallel form of the posttest, e.g., items on the two were considered matched in complexity. This test was administered three weeks after the experiment to compare performance of

the forward, reverse, and random sequence groups in retention of the final task I,1.

Attitude Measure

The rationale for including an attitude measure, entitled "Learning Program Questionnaire," was that Ss with the forward sequence would be expected to have a more favorable attitude toward the program than would Ss with either a reverse or randomly sequenced program. The attitude measure was administered during the experiment, after each S had completed the assigned one of the three sequence versions of the program, and prior to administration of the posttest. A portion of the instructions for the questionnaire explained to Ss that the results of the questionnaire would not affect how well they did on the learning program. The maximum favorable attitude score was 100, or from 0 to 20 possible points for each of the five attitude statements.

The questionnaire was developed according to recommended techniques of attitude scale construction. There were five descriptive statements about the learning program, which were identified as a questionnaire rather than an attitude measure, for the purpose of obtaining an objective evaluation of Ss attitude toward the program. Ss were to respond to the five statements along a five-point continuum: "strongly agree," "agree," "undecided," "disagree," and "strongly disagree;" so that responses were likely to cover at least a three-point positive or negative attitude range

about the program, since respondents to attitude scales tend to use those ranges within the two extremes. Two precautions were taken in the attempt to correct for such effects as the "halo effect;" first, Ss were instructed not to respond according to their overall impression, but instead to consider each statement separately, and second, the direction (left to right location of a + and - series of reactions) of favorable or unfavorable judgement was not the same for all the attitude statements. The "Learning Program Questionnaire," can be found in Appendix G.

Conduct of the Experiment

The experiment compared the learning program with instructional units sequenced in the order implied by task I,1 of the hierarchy (forward sequence), with reverse and random sequence presentations.

Since there is evidence that, as they study, Ss of sufficient mental age can internally or implicitly resequence and reorganize material for themselves (see Chapter II), a treatment by levels analyses of variance design was used to evaluate the treatment effect (sequence variations) upon performance, as well as the interaction of mental ability and sequence upon performance. Figure 3 presents the research design used for the experiment.

Fig. 3.--Research Design

		Sequence Presentation		
		Learning Hierarchy	Reverse	Random
Mental Ability	High			
	Average			
	Low			

The hypotheses to be tested were:

1. The time to completion will be greater for the reverse and randomly sequenced programs, than for the program sequenced according to the learning hierarchy.

2. Program error rates during instruction will be greater for the reverse and randomly sequenced programs, than for the program sequenced according to the learning hierarchy.

3. Competency test results will be greater for the program sequenced according to the learning hierarchy, than for reverse and randomly sequenced programs.

4. Attitude toward the learning program will be more favorable for the program sequenced according to the learning hierarchy, than for the reverse and randomly sequenced programs.

5. Posttest performance will be greater for the program sequenced according to the learning hierarchy, than for reverse and randomly sequenced programs.

6. Retention, measured after a period of three weeks, will be greater for the program sequenced according to the learning hierarchy, than for the reverse and randomly sequenced programs.

7. Interaction is expected between sequence presentation and mental age level on achievement.

Individuals were randomly assigned by stratified I.Q. level to the three sequence presentations (treatments) in this manner. Names of the 60 Ss with I.Q. 80 and above, in the four classrooms used for the experiment, were listed by I.Q. scores, from low to high. (Names of Ss having the same I.Q. score were alphabetized.) Three pieces of paper, on each of which was written the number of a sequence group (forward - 1, reverse = 2, random = 3), were placed in a box. The first piece of paper randomly chosen from the box was numbered 2, so the first of two names at I.Q. score 80 was assigned to treatment 2 (reverse sequence). The next name was assigned to treatment 3 (random sequence), the third name to treatment 1 (forward sequence), the fourth name to treatment 2 (reverse sequence), and so on down the list. There were 20 Ss in each of the three treatment groups. The I.Q. score ranges of the mental ability levels for the experiment were: high, 105-126; average, 92-104; and low, 80-91.

Directions to Subjects

Two pages of directions for taking the program on a completely self-instructional basis were read aloud to the

Ss, as they read them to themselves at their desks. These instructions make up the first two pages of the learning program (see Appendix E). As part of these instructions, Ss were told not to change an answer and not to skip or to go backward in the program. Ss were encouraged to do their best, and a prepared "motivational statement" was read informing them that they would receive 100 points to buy classroom favors for taking the program, and that their classroom teacher would see the results of the test they took after completing the program.

Ss were told that they all had the same learning program, but that the program had been put together in three different ways, so that the order of the instruction varied. It was explained that a list of their names had been used to randomly assign them to one of three types of instructional sequence. Further, they were told that one program sequence "teaches algebra in the order we feel it should be taught," another program sequence "is the exact opposite of this, that is, it starts out with the most difficult material first and gets progressively easier," and a third sequence "has a mixed-up order, with difficult and easy parts alternating throughout."

Procedures

Administration of the learning program, attitude scale, and tests was carried out in each classroom by the teacher and E.

Each S received a learning program coded to his experimental condition, with his name written on it. The experimental code 0 identified 48 Ss who were not included in the experiment because of (a) below 80 I.Q. score, (b) no record of I.Q. test administration, or (c) a missing cumulative folder. Continuous monitoring was necessary during administration of the program so that Ss would not (a) look at the correct answer until they had made a response, (b) change an incorrect answer, or (c) skip or go backward in the program.

As Ss finished the learning program, the date and time was noted on the front of their program booklets. Then, after Ss completed the attitude scale, they received a posttest.

The majority of Ss finished the program in a five-day period, or by the end of the school week. The retention test was administered three weeks later.

The learning program[2]. The forward sequence group received the program which presented the instruction in the sequence implied by the hierarchy. The reverse sequence program began at instruction unit 10, with 1 last. The random sequence, determined by randomly drawing slips of paper numbered 1 to 10 from a box, was: 8, 4, 9, 5, 10, 7, 6, 1, 3, 2.

---

[2]The Basic Algebra Skills Learning Program may be found in Appendix E.

The program format was so planned that the program only had to be reproduced once, and could then be put together in either of the three instructional sequences. Competencies were identified by roman numerals, and frames within competency levels by arabic numbers. Thus, by rearranging pages in the appropriate order, the program could be assembled to teach the competency levels of task I,1 of the hierarchy in a forward, reverse or random order.

The program was written on standard 8 1/2" by 11" paper. The left half of the page was instruction, and answers to the blanks were on the right-half of the page. The answer for each frame was written in the frame below it. Ss used the shield provided to cover the right side of the page as they worked through the program, and as a marker to keep their place in the program from day to day.

Test procedures.The classroom teacher and E proctored the administration of the tests and attitude scale. Also, Ss were assigned additional work at their desks, so that those who finished before their classmates would not disturb those still working with the experimental materials.

The competency test, the last frame of each instructional unit of the program, was identified in the program as a "Test Question". Ss were instructed prior to taking the program to respond to the Test Question frames in the same way they would the other frames in the program. That is, to "read the frame, and write an answer in the blank. Then

slide the shield down to see that your answer is correct" (see the directions for taking the program in Appendix E).

Ss were handed the attitude scale as they finished the program and told to ask any questions they might have before answering the questionnaire.

The Ss also understood prior to their taking the program that they would receive a test (posttest) after completing the program. When Ss were given the posttest, they were told that this was a test over what they had learned from the program. Later during administration of the retention test, it was explained to Ss that the test was being administered to determine how much they remembered of what they learned from the program.

CHAPTER IV

RESULTS

This study involved three phases of work: (a) the development of a program seqeunced in accord with competency levels of a learning hierarchy; (b) successive tryouts and revisions of the program so sequenced, in order to improve the effectiveness of the instruction prior to its use in the experiment; and (c) an experiment comparing this hierarchy (forward) sequence with reverse and random program sequences.

The previous chapter describes (a), and gives the detailed procedures involved in (b) and (c). This chapter will present the information from the pretest administration, and the results of steps (b) and (c) in two distinct parts. First, the results of the successive tryouts of the program will be given, based on individual tryouts and small group tryouts (formative evaluations), and a third and final tryout with a larger group (summative evaluation). Second, the experimental results comparing forward sequence with reverse and random sequences, will be reported.

Performance Pretest

The 10-item pretest for task I,1 of the hierarchy, was administered to all five classrooms involved in the study.

In the classroom used for formative revisions and summative evaluation of the program only a single S obtained a score of 1. In the four classrooms used for the experiment, five Ss received a score of 1. None of the other Ss in these five classrooms answered any items correctly on the pretest. This is assurance that these Ss had not already mastered the final task. In fact, since so few Ss could answer even one question on the pretest, the posttest scores can practically be considered as gain scores.

Formative Evaluations and Summative Evaluation of the Learning Program

The results of the three tryouts of the program will be presented in this section. The individual tryouts and small group tryouts are called formative evaluations, because the data derived from them were subsequently used to revise the program. (These program revisions were reported in Chapter III.) The third and final evaluation of the program on another larger group, the summative evaluation, was used to estimate the teaching effectiveness of the program prior to its use in the experiment.

Four criteria were used to evaluate program effectiveness during the small group formative evaluation and for the summative evaluation. These were: time to complete the program, program frame error rate, competency test score, and posttest score. Since the purpose of individual tryouts of the program was to uncover program inadequacies by having

the student answer aloud, and to question ambiguous portions of the program, no record of program errors was kept nor a posttest administered. However, the time taken by each S to complete the program, and the score on the competency test were used as criteria for the individual tryout formative evaluation.

Formative Evaluations

The individual tryout formative evaluation was made after the first revision of the learning program. The small group formative evaluation was made following the second revision of the learning program. The third and final evaluation, used as a summative evaluation, was made after the third revision of the program. The results of these evaluations, conducted after each of the three revisions of the learning program, are presented and compared below.

Time-to-completion. Three Ss, one from each of high, average, and low mental ability levels, took an average of 5 hours and 8 minutes to work through the program individually with E. Another three Ss who took the revised program in a small group situation, completed the program in an average of 3 hours and 30 minutes. The nine other Ss who participated in the final evaluation of the learning program, finished an average of 10 minutes sooner than had the small group, or in 3 hours and 20 minutes. Program revisions thus resulted in a reduction of 1 hour and 48 minutes to complete the program. Some of this reduction in the time to complete

the program between individual tryouts and the final evaluation may be accountable to the time Ss spent during the individual tryouts responding aloud, commenting, and questioning ambiguous portions of the program.

Program errors. The program errors were also substantially reduced through successive revisions of the learning program. Ss who received the small group tryout of the program had an average of 54.33 program frame errors. However, in the final evaluation, mean program errors was lowered to 31.44. Therefore, there was an average reduction of 22.89 program errors.

Competency test. There was 1 test item per competency level for task I,1 of the hierarchy. The mean score for the competency test in individual tryouts was 5.33. A second program revision resulted in a small group average competency test score of 6.33. Following further program revision, the final evaluation group mean increased to 7.89, out of a possible maximum of 10. Thus, successive program revisions produced an average competency test score increase of 2.56, or a gain of two and one-half competencies. Prior to the final evaluation the competency test items were included in the program as the last regular teaching frame of each instructional unit. However, for the final evaluation, each competency test item was identified as a "Test Question." This format change may have contributed to the competency test score gain on the final evaluation.

Performance posttest. The 10-item posttest for the final task I,1 of the hierarchy, resulted in a small group mean of 5.67, and a final evaluation mean of 4.89, or a decrease of .78 of a score. Two of the nine Ss who participated in the final evaluation of the learning program received a score of 0 on the posttest. (The experimental results of the forward sequence group in the experiment might constitute a better summative evaluation since more Ss were involved, and the program was administered to entire intact classrooms, rather than to just those Ss who had not participated in earlier evaluations of the learning program in the classroom used for formative revisions).

Summary of formative evaluation results. The results of the formative evaluations indicated that the revisions of the program succeeded in reducing the time to complete the program, lowered the program frame error rate, and increased the number of competencies gained. There was no increase of average posttest score, but rather a decrease of .78 of a score, between the small group and final evaluation means.

Program revisions resulted in a reduction of 1 hour and 48 minutes to complete the program, or a 35% time reduction. The program errors were lowered by 42%, or an average of 22.89 program errors. A mean two and one-half competencies were gained, or a 48% increase.

Summative Evaluations

During the summative evaluation the average time taken to complete the program was 3 hours and 20 minutes, or within

approximately four and one-third classroom periods. The program had an acceptable average frame error rate of 15%. The summative evaluation goal was to establish program effectiveness for each competency level and task I,1 to the 85/85 criterion (85% of the learners will score 85% or higher). However, this final evaluation resulted in a competency test standard of 85/79, 85% of the learners scored 79% or higher. The posttest results were 67/49, or 67% of the learners scored 49% or higher.

Although the 85/85 criterion was not reached, successive formative evaluations had provided evidence of increased program effectiveness, and the time for formative revisions of the program had elapsed; it was therefore a matter of economy to accept a somewhat lower standard. It may also be recalled that the population of the tryout school may be considered a disadvantaged population. It is suspected that in a more advantaged school, the design criterion of 85/85 may have been met or surpassed.[1]

Experimental Comparison of Hierarchy, Reverse, and Random Sequence Groups

Six treatment by level analyses of variance were made in comparing the forward, reverse, and random groups by mental ability levels. The dependent variables were: (a) time to complete the instruction; (b) program error rate; (c) competency test score; (d) attitude measure; (e) posttest score; and

[1]The list of scores on successive tryouts of the program can be found in Appendix H.

(f) retention test score. The results of the sequence variations, as well as the interaction of mental ability and sequence, are reported.

Analysis of these six dependent variables on the regular Biomedical 05V computer program yielded nonsignificant results. A study of the variability revealed it to be heterogeneous rather than homogeneous.[2] Tables 1-6 present the raw score means and standard deviations for the six dependent variables for the regular Biomedical 05V computer analyses. Evidences of heterogeneous variability were: instances when the variability was greater than the mean; unequal variance among the low ability groups, and between the high and low ability random groups; and the extreme variation of the low random group. Therefore, a square root score transformation analysis was performed. This transformed score is the square root of the number, plus the square root of the number plus one ($\sqrt{x} + \sqrt{x + 1}$). The following are the results of the six square root score transformation analyses of variance.

Time to Complete The Instruction

Table 7 presents the means and standard deviations for time to complete the instruction for the forward, reverse, and random groups by ability levels. Though the forward

---

[2]The raw scores for I.Q. and for the six dependent variables are listed by sequence group and mental ability level in Appendix I.

TABLE 1

RAW SCORE MEANS AND STANDARD DEVIATIONS: TIME TO COMPLETE THE INSTRUCTION

Group	Ability Level	Mean ($\overline{X}$)	S.D.	N
	High	177.86	51.95	
Forward	Average	193.57	46.16	20
	Low	211.67	45.46	
	High	185.83	51.32	
Reverse	Average	220.00	49.16	20
	Low	231.43	77.17	
	High	217.50	54.75	
Random	Average	203.57	39.97	20
	Low	225.71	102.45	

TABLE 2

RAW SCORE MEANS AND STANDARD DEVIATIONS: PROGRAM ERROR RATE

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	15.00	19.03	
Forward	Average	28.43	17.73	20
	Low	28.83	24.18	
	High	19.17	21.40	
Reverse	Average	30.57	22.19	20
	Low	46.00	35.85	
	High	23.17	4.53	
Random	Average	40.86	12.80	20
	Low	67.00	40.32	

TABLE 3

RAW SCORE MEANS AND STANDARD DEVIATIONS:
COMPETENCY TEST

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	9.14	.69	
Forward	Average	7.71	1.89	20
	Low	8.17	1.47	
	High	8.67	1.50	
Reverse	Average	7.71	1.70	20
	Low	7.00	2.58	
	High	8.50	1.52	
Random	Average	7.71	1.11	20
	Low	7.00	2.31	

TABLE 4

RAW SCORE MEANS AND STANDARD DEVIATIONS:
ATTITUDE MEASURE

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	77.14	13.18	
Forward	Average	58.57	9.88	20
	Low	67.50	8.22	
	High	76.67	16.02	
Reverse	Average	60.00	15.27	20
	Low	61.43	27.19	
	High	70.83	8.61	
Random	Average	57.86	10.35	20
	Low	70.71	18.58	

TABLE 5

RAW SCORE MEANS AND STANDARD DEVIATIONS: POSTTEST

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	5.43	3.41	
Forward	Average	4.29	1.11	20
	Low	5.33	3.14	
	High	6.33	3.61	
Reverse	Average	3.71	2.36	20
	Low	3.71	2.87	
	High	6.33	2.16	
Random	Average	3.71	1.60	20
	Low	2.57	3.10	

TABLE 6

RAW SCORE MEANS AND STANDARD DEVIATIONS: RETENTION TEST

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
Forward	High	5.43	3.91	20
	Average	1.86	1.77	
	Low	6.17	3.82	
Reverse	High	7.00	4.10	20
	Average	3.71	3.73	
	Low	2.71	3.50	
Random	High	6.83	2.48	20
	Average	4.29	3.30	
	Low	3.29	3.35	

group overall and by ability levels, took less time to complete the learning program, there were no statistically significant differences (see Table 13 of F ratios).

Program Error Rate

The forward group had significantly fewer program errors than the random group, but there was no significant differences between the forward and reverse groups (Table 8). These combined group transformed score means are: forward group, 9:00; reverse group, 10.28; and random group, 12.71. There was no significant interaction effect (Table 13) between sequence variations and mental ability levels on program error rates.

Competency Test Score

Table 13 indicates there were no significant differences among treatment groups, or for interaction between sequence presentation and mental ability levels on the competency test. Nor was there a trend in favor of one of the sequence groups: the forward, reverse, and random groups performed about the same on the competency test (Table 9).

Attitude Measure

Ss in the forward group rated their program higher than did Ss in the reverse and random groups (Table 10). However, as Table 13 shows, there were no statistically significant differences either among the treatment groups, or for interaction of sequence variations and mental ability level upon student attitude toward the program.

TABLE 7

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES: TIME TO COMPLETE THE INSTRUCTION

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	26.49	3.69	
Forward	Average	27.69	3.32	20
	Low	28.99	3.16	
	High	27.07	3.90	
Reverse	Average	29.56	3.14	20
	Low	30.13	4.82	
	High	29.34	3.65	
Random	Average	28.46	2.71	20
	Low	29.52	6.26	

TABLE 8

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES:
PROGRAM ERROR RATE

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	6.66	4.51	
Forward	Average	10.19	3.73	20
	Low	10.14	4.17	
	High	7.62	4.94	
Reverse	Average	10.38	4.39	20
	Low	12.83	4.99	
	High	9.69	.96	
Random	Average	12.70	2.16	20
	Low	15.74	5.09	

Mean difference value between the Forward and Random Groups	$t = 2.90^{**}$ $df = 51$

$^{**}p < .01$

TABLE 9

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES: COMPETENCY TEST

Group	Ability Level	Mean ($\bar{X}$)	S.D.	N
	High	6.20	.22	
Forward	Average	5.70	.67	20
	Low	5.87	.51	
	High	6.03	.50	
Reverse	Average	5.70	.61	20
	Low	5.40	.98	
	High	5.98	.52	
Random	Average	5.72	.39	20
	Low	5.41	.88	

TABLE 10

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES: ATTITUDE MEASURE

Group	Ability Level	Mean ($\overline{X}$)	S.D.	N
	High	17.57	1.53	
Forward	Average	15.32	1.34	20
	Low	16.47	.98	
	High	17.47	2.00	
Reverse	Average	15.45	1.94	20
	Low	15.33	3.84	
	High	16.86	1.05	
Random	Average	15.22	1.40	20
	Low	16.74	2.28	

Posttest Score

Table 11 shows the forward group did only slightly better on the posttest than the reverse and random groups. It also shows that the average ability and low ability forward groups performed better than the average and low ability reverse and random groups. However, Table 13 indicates no significant treatment differences or significant interaction.

Retention Test Score

There were no significant differences among the treatment groups for retention. The random group mean was superior to the forward and reverse group means (Table 12). Also, there was no significant interaction effect between sequence presentation and mental ability levels on retention. However, it is interesting to note, that the highest F ratio for interaction in Table 13 occurs for the retention test. A study of the means in Table 12 reveals that the low forward group not only performed better than either the low reverse and low random groups, but was superior to the high forward group on the retention test. This may be seen even more clearly from the list of raw score means in Appendix I.

Ability Levels Across Treatments on the Dependent Variables

Table 13 presents the analyses of variance F ratios for the sequence variations, by ability level groups, and for sequence by ability interaction. The results have been

TABLE 11

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES: POSTTEST

Group	Ability Level	Mean ($\overline{X}$)	S.D.	N
	High	4.67	1.45	
Forward	Average	4.34	.51	20
	Low	4.65	1.41	
	High	4.87	2.01	
Reverse	Average	3.85	1.48	20
	Low	3.77	1.68	
	High	5.16	.87	
Random	Average	4.04	.73	20
	Low	3.03	1.80	

TABLE 12

MEANS AND STANDARD DEVIATIONS OF TRANSFORMED SCORES: RETENTION TEST

Group	Ability Level	Mean ($\overline{X}$)	S.D.	N
	High	4.50	1.96	
Forward	Average	2.63	1.54	20
	Low	4.79	2.05	
	High	5.09	2.17	
Reverse	Average	3.57	2.11	20
	Low	2.94	2.10	
	High	5.35	.93	
Random	Average	4.04	1.74	20
	Low	3.50	1.76	

reported of the sequence variations and for interaction of sequence and ability, for the six dependent variables. This section presents the combined group means for the significant differences among high, average, and low mental ability level groups, across the sequence variations, for the six dependent variables.

Of the six dependent variables, there were four significant differences for mental ability level groups (Table 13): program errors ($p < .01$); competency test score ($p < .05$); attitude measure ($p < .05$); and retention test score ($p < .05$). The combined group, transformed score, mental ability level means for these four dependent varialbes, are as follows:

1. program errors; high ability, 7.99; average ability, 11.09; and low ability, 12.90;

2. competency test score; high ability, 6.07; average ability, 5.70; and low ability, 5.56;

3. attitude measure: high ability, 17.30; average ability, 15.33; and low ability, 16.18;

4. retention test score: high ability, 4.98; average ability, 3.41; and low ability, 3.74.

Summary of Results

This Chapter presented the information from the pretest administration, the results of the formative evaluations and the summative evaluation of the learning program, and the results of the experiment which compared the

TABLE 13

F RATIOS FOR TRANSFORMED SCORE ANALYSES OF VARIANCE: SEQUENCE VARIATIONS, ABILITY LEVELS, AND SEQUENCE BY ABILITY INTERACTION

Source of Variation	df	Time to Completion	Program Errors	Competency Test Score	Attitude Measure	Posttest Score	Retention Test Score
Sequence Variations	2	.68	4.08*	.75	.16	.64	.29
Ability Levels	2	1.08	6.87**	3.30*	4.62*	3.08	3.77*
Sequence by Ability Interaction	4	.33	.27	.30	.48	.92	1.29

$*p < .05$
$**p < .01$

hierarchy (forward) program sequence with reverse and random sequences.

Pretest administration in the five classrooms used for the study resulted in only six Ss receiving a score of 1, of a possible 10, correct.

The formative evaluation results showed that the three successive revisions of the program produced a 35% reduction in average time to complete the program, lowered program frame errors by 42%, with a 48% gain in the number of competencies mastered. There was a decrease of .78 of a score between the small group formative evaluation posttest mean, and the final evaluation (summative evaluation) posttest mean, in which two of the nine Ss in the final evaluation received 0 scores on the posttest.

For the summative evaluation, Ss took an average of 3 hours and 20 minutes to complete the program. The program frame error rate was 15%. Competency test results reached 85/79 (85% of the learners scored 79% or higher), and posttest results were 67/49 (67% of the learners scored 49% or higher).

Experimental comparison of the forward, reverse, and random sequence groups by mental ability level was performed in six treatment by level analyses of variance. Analysis of the dependent variables on regular Biomedical 05V computer program resulted in nonsignificant results, and revealed the variability to be heterogeneous rather than homogeneous. Therefore, a square root score transformation analyses of

variance was performed. Among the six score transformation analyses of variance, there was one significant difference, that of fewer program errors in favor of the forward group. There were no other significant differences for sequence variations. There were four significant differences for mental ability level groups across dependent variables: program errors; competency test score; attitude measure; and retention. No significant interaction effects were found for sequence and mental ability upon performance.

However, on three of the other five dependent variables, the forward group mean was higher than the reverse or random group means. That is, the forward group took less time to complete the program, rated the program higher, and scored slightly better on the posttest. All three sequence groups performed about the same on the competency test. The random group mean performance was better than the forward group mean for retention. Also, though there was no significant interaction between sequence presentation and mental ability levels on retention, the highest interaction F ratio occurred on the retention test, in which the low forward group was superior to the high forward group on retention.

CHAPTER V

CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS FOR FUTURE RESEARCH

Development Of The Experimental Program

It was pointed out in previous chapters that the original program used in this study, the Princeton Algebra Program, existed prior to the learning hierarchy. Gagné and Paradise had used the existing program to define a hierarchy for the task solving linear algebraic equations. This section discusses the experience of sequencing the instructional units of the program in the order implied by the competencies of the hierarchy.

The initial classifying of individual program frames to match the competencies of the hierarchy, was a long and painstaking job. After the entire program had been carefully studied, each of the 232 frames was matched to one of the 22 hierarchy competencies. Even though only the task I,1 branch of the hierarchy was used in this study, all of the Princeton Algebra Program frames were classified, because it was desired to ensure the inclusion of all the instruction needed to teach each competency, and also that there be no extra frames which would provide for review or teaching ahead over competencies.

It was difficult to accept the initial classification of frames which revealed either a scarcity of frames, or no frames at all, for some of the hierarchy competencies. Consequently, each frame classification was re-studied, only to produce the same end results. Since instructional materials are often merely assumed to teach for all the competencies required for learning, the identification of competencies needed to perform a task, and a careful matching of instruction to these competencies, should help to ensure the completeness of the instruction and thus facilitate learning.

Since the frame sort had revealed that the original program had no teaching frames, or few teaching frames, for several of the competencies, extensive writing became necessary before the program could be tried out the first time. The result was a drastically revised program, with 136 new frames added to 69 frames from the original program.

In an effort to use every frame that matched a competency, the lower-level competencies had an excess of frames, and thus lower instructional units of the program were too lengthy and overtaught. On the other hand, the higher-level competencies did not have sufficient practice frames and the instruction steps taken were too large. Therefore, in addition to rewriting for clarification, later revisions deleted frames from lower-level program instructional units, and added to higher-level instructional units. These deletions of frames at lower level instructional units resulted in use of only 33 frames from the original program in the new Basic Algebra Skills Learning Program.

For the purposes of the experiment, each instructional unit of the program had to be complete in itself, that is, to provide as good instruction as possible, but there could be no review of earlier competencies or teaching ahead to later competencies of the hierarchy. Not only is this condition difficult to provide for meaningful curriculum material, while it is theoretically possible, it is practically nearly impossible. The best way to explain this, is to give an example. At a lower-level competency, the student learns that $x = 1x$. Then later, at a higher-level competency, he learns to combine like terms, such as $3x + x = 4x$. Now, theoretically if the student has not learned the lower competency $x = 1x$, he cannot perform $3x + x = 4x$. However, in practice by giving the student examples, and guidance in leading him toward a solution at the higher-level competency, it is possible he may learn the lower-level competency at the higher-level, without first having learned it at the lower level. Thus the experimental requirement for testing the hypothesis underlying this study somewhat hampered the writer's practical programming tendencies.

Formative Evaluations

The purpose of the formative evaluations (program tryouts) was to uncover and correct program inadequacies, for the best first efforts of the programmer may prove to be totally ineffective with the learners. An instructional point that is crystal clear to the programmer may completely

elude the understanding of the learner. Three phases of the formative evaluations were most helpful to the programmer in revising the program.

Of all three tryouts of the program, using individual Ss, a small group of Ss, and another larger group of Ss, the individual tryouts of the program proved to be the most useful in making revisions of the program. The junior high school Ss used for the individual tryouts were not at all reticent in making comments and questioning those portions of the program that were ambiguous to them. Students just do not think as the programmer imagines they will, and this phase of the program tryouts enables the programmer to observe first-hand just how close he came to anticipating and preparing for Ss' instructional needs.

The small group tryout aided the programmer in two ways. First, it helped in preparing the directions for administration of the program on a completely self-instructional basis. Before the program was administered to the small group of Ss, they were given what was considered to be complete and basic directions for taking the program. Then, each question that Ss asked after they had been given the directions was made note of and the answer to each question was included as part of the final directions for taking the program on an individual basis. As it turned out, these directions proved to be quite adequate for enabling Ss to take the program on a completely self-instructional basis, both during the summative evaluation, and later for the

experiment. And second, particularly helpful about the small group tryouts, were the interviews held with individual Ss following their completion of the program. At that time, each student indicated areas of the program that "talked down to him," were too difficult, or were not clear to him. The data from the interviews were specifically used to clarify and simplify the size of instructional steps in the program.

Summative Evaluation

The results of the summative evaluation were used as an indication of the effectiveness of the program, prior to its use in the experiment using different sequencing three groups. This section describes the difficulties encountered during the summative evaluation of the learning program.

Only 9 Ss were used in the summative evaluation, for they were the remaining Ss with I.Q. 80 and above, in the classroom used for formative evaluations and summative evaluations, after using three Ss for individual tryouts and another three Ss for small group tryout of the program. (The overall I.Q. mean was lower for the summative evaluation group than the I.Q. mean for either the individual tryouts or small group tryouts, as can be seen in Appendix H.)

The summative evaluation was held the week prior to Christmas. A surprise assembly called the second day necessitated that slower Ss, and those who had been absent, complete the program after the Christmas vacation. This circumstance, plus the fact that preparation of the experimental program required there be no review or teaching ahead

over competencies, and a demanding posttest, made reaching the summative evaluation design criterion even more difficult.

It has been speculated (Chapter III) that the forward group in the experiment probably could be viewed as a better summative evaluation group, and that they would perform better on the posttest than the actual summative evaluation group. As it turned out, the forward group did perform slightly better than the summative evaluation group on the posttest.

Since this school may be characterized as having a large proportion of disadvantaged pupils, it may be that with a summative evaluation of the program in a more advantaged school, the design criterion of 85/85 (85% of the learners score 85% or higher on a posttest) could be met.

Experimental Results

This section will review the results of the experiment comparing the hierarchy (forward) program sequence with reverse and random sequences, and draw conclusions from these results.

First, it will be recalled that the variability within sequence groups made it difficult to detect significant differences among the sequence groups. Thus, a transformed score analyses of variance was performed.

The transformed score analyses of variance revealed one significant difference due to treatment effect, that of program errors, in favor of the forward group. There were

no other significant differences for treatment variations, although forward group means were superior to reverse and random group means for the dependent variables: time to complete the instruction, attitude measure, and posttest score. All three sequence groups performed about the same on the competency test. The random group mean was superior to forward and reverse group means on the retention test, as though the effort in mastering successive competencies made them highly resistant to forgetting.

There were no significant interaction effects for sequence and mental ability level on performance. However, the low ability forward group took less time to complete the program, had fewer program errors, scored higher on the competency test, scored higher on the posttest, and had better retention, than either the low reverse or low random groups. Also, the low forward group performed about the same on the posttest as did the high forward group and had superior retention.

Thus, this is seen as evidence that the forward sequence does help low ability Ss. Further research may reveal that lower ability Ss receive greater instructional benefit from a careful instructional sequence than do high or average ability Ss who are better able to resequence and reorganize material for themselves.

Comparison Of This Study With Other Research

The research of Gagné and his associates is involved with testing the predictions of transfer among subordinate learning sets (competencies) of a learning hierarchy supporting a final task. These studies are not concerned with an instructional sequence based on the competencies of the hierarchy. Perhaps the best way to contrast the studies of Gagné and others with this study, is to compare the manner in which Gagné and Paradise utilized the Princeton Algebra Program, with the way that program was used in this study. After Gagné and Paradise had administered the program, (which was not sequenced according to a hierarchy), Ss received a test over the competencies in the hierarchy. The prediction of transfer among competencies was studied by noting the pass-fail patterns between competency test items.

This study first sequenced the program according to a learning hierarchy, and compared the program so sequenced with reverse and random program sequences. This was a study involved with the applicability of a hierarchy for the sequencing of instruction.

The majority of what are termed sequencing studies in the literature are experiments comparing logical versus scrambled frame sequences. They do not attempt to apply a learning structure, but compare logical and scrambled program frames. This study sought to determine how learning is affected by an instructional sequence derived from, and

based on, the transfer implications underlying a task analysis and learning hierarchy. To do this, a learning hierarchy was used to develop an instructional sequence using actual curriculum material, and program presentation of instructional units were varied, rather than program frames.

Suggestions For Future Research

Considering the one significant difference and other trends in favor of the forward sequence, and evidence that low ability Ss benefit more than do high or average ability Ss from this sequence based on the competencies of a hierarchy, continued research related to this study is believed worthwhile to further study the implications of a task analysis and learning hierarchy for the sequencing of instruction. These research recommendations are briefly summarized as follows:

1. There is a need for intermediate research between the laboratory-type studies of Gagné and his associates, and the overall sequencing study in an instructional situation reported in this dissertation. An example of such research would be an experiment in which the competency test be presented as it was in this study, that is, included within the program as the last frame of each instructional unit. However, analyses would consist of the forward, reverse, and random groups' degree of mastery at successively higher competency levels, rather than only consideration of total competency test score.

2. The instructional sequence of the competencies of task I,1 of the hierarchy could be varied for study. For example, there are two sub-branches, under the task I,1 branch of the hierarchy. Rather than sequencing horizontally across a level, before proceeding to the next level, one of the sub-branches could be completely taught vertically, before teaching the second sub-branch.

3. Chapter III of this study reported the decision had to be made following individual tryouts of the program as to whether to provide additional instruction at three higher competency levels in the hierarchy, or to add competencies to the hierarchy. It will be recalled that for the purposes of this study the former course was chosen. However, an important contribution could be made toward the study of instructional sequencing by deleting or adding competencies to the hierarchy, administration of learning programs whose instructional units are in accord with deletions and additions of competencies, and then testing for final task performance.

4. This Chapter earlier discussed the possibility that <u>S</u>s in the reverse and random groups may have learned a lower-level competency at a higher-level competency without first having learned it at the lower-level competency, by being provided with examples and practice at the higher-level competency. An experiment to determine if lower-level learning may be obtained a. higher-level learning through practice, would be one in which one group would receive the

entire learning program and another group only the upper half of the program, followed by comparison of the two groups on final task performance.

5. There is a need for research as to whether transfer, the recall of previous relevant learning and use of this learning during the learning of something new, is facilitated to a greater degree by instruction with a program sequenced according to a learning hierarchy, than by reverse and randomly sequenced programs.

6. Use of the entire hierarchy rather than only one branch of the hierarchy to sequence a learning program would result in instruction of longer duration, which when presented in forward, reverse, and random sequences may produce more pronounced differences in performance among high, average, and low mental ability levels.

7. There is a need for studies to validate the structure of learning hierarchies. A technique for doing so was reported by Resnick and Wang (1969). The successful validation procedure was based on the examination of pass-fail contingency tables for all possible pairs of items in the hierarchy. Phi/Phimax coefficients were computed for each table. When the coefficient was at or above some arbitrarily defined level, a hierarchical relationship between the two items was inferred. On the basis of these simple prerequisite relationships, it was possible to construct a hierarchy which could have both linear and branching sections.

Follow-Up Study

A follow-up study to this dissertation will be conducted with support by the Office of Education. This study will compare the forward sequence of the Basic Algebra Skills Learning Program with a random sequence using a larger number of Ss (approximately 150 per group), preferably 9th graders rather than 8th graders, in a school with more normal distribution of I.Q. scores. The forward version of this two-group sequencing experiment will constitute another summative evaluation of the learning program. Also, the program errors made by the larger number of Ss in this follow-up study could be used in a path analysis technique described by Spady and Greenwood (1969) to empirically determine the most appropriate instructional sequence for the competency levels of task I,1 of the hierarchy.

APPENDIX

APPENDIX B

SUBORDINATE COMPETENCIES OF TASK I,1 OF THE SOLVING EQUATIONS LEARNING HIERARCHY

Hierarchy Box	Competency Number	Description	Test Item
IVA,1	1	Equivalence of 1x and x	Which term below is the same as x? (sym) ________ 3x; 10x; -2x; y; 1x; 8x; -y
IVA,2	2	Identification of an equation (= sign)	Which of the following is an equation? (sym) ____________ $8y + 2 - 10$ $(8y + 2)10$ $8y + 2 = 10$ $\frac{8y + 2}{10}$
IV,2	3	Recognition of equivalent terms	1. 3(a-b) 1. ________ 2. 12n 2. ________ 3. 6ab 3. ________ 4. $\frac{5}{y}$ 4. ________ 5. 16 5. ________ For each of the above terms, there is one term below which can be combined with it. Find the term below which can be combined with 3(a-b), then the term which can be combined with 12n, and so on. 14s 9q 5m -6n 5p 40 -4a 1pr -12ab $\frac{11}{x}$ (a-b) $\frac{2}{y}$ 16a (x-y)

Hierarchy Box	Competency Number	Description	Test Item
IV,3	4	Performing addition and subtraction of numbers in sequence	$3 + 9 + 27 - 13 + 2 - 39 - 5 + 68 =$ (sym) ____
IIIA,1	5	Combining fractions with like denominators	$\frac{3}{2x+y} + \frac{4}{2x+y} - \frac{2}{2x+y} =$ (sym) ____
III,2	6	Addition and subtraction of terms in sequence	$2y + 7z - 8y - 4w + y + 14w - z =$ (sym) ____
III,3	7	Supplying sum and difference equivalents to sums and differences (arith. numbers)	$2 + 7 - 5 + 1 - 6 = 3 - 6 + 8 + 2$ ($\pm$?) (sym) ____
II,1	8	Supplying sum and difference equivalents to sums and differences (terms)	$3g + 5f + h = 2g - 3f - 4h$ ($\pm$?) (sym) ____

Hierarchy Box	Competency Number	Description	Test Item
II,2	9	Simplifying an equation by adding and subtracting arith. numbers to both sides	Solve for <u>w</u>: $16 + w - 10 = 7 + 4 - 2$ $w =$ (#) ________
I,1	10	Simplifying an equation by adding and subtracting terms to both sides	Solve for <u>b</u>: $7b + 2a + 3b - a = 10b + 2a + 35 - b$ (sym) ________

APPENDIX C

PRETEST AND POSTTEST FOR

THE FINAL TASK I,1

SIMPLIFYING EQUATIONS

Date ______________________ Name ____________________

1. Solve for x: $10x + 4 = 9x + 6$

2. Solve for y: $9y + 3z = 8y + 6z$

3. Solve for e: $5e + 6f = 12f + 4e - 2f$

Name ____________________

4. Solve for g: $4g + 4h = 2g + 9h + g$

5. Solve for c: $6c - 2d = 8d - 4c - 7d + 9c$

6. Solve for a: $8a + 3b - 5a = 7b + 2a + 2b$

7. Solve for v: $2v + 2w + 3w = 7w + v$

Form P II

Name ________________

8. Solve for s: $7s + 2t - s = 8t + 5s + 20$

9. Solve for n: $5n + 3p + 4n - p = 8n + 6p + 18 - p$

10. Solve for b: $7b + 2a + 3b - a = 10b + 2a + 35 - b$

APPENDIX D

EXPLANATION FOR AND DESCRIPTION OF ADDITIONAL INSTRUCTION PROVIDED FOR COMPETENCIES 6, 10, AND 7

Competency 6 description and test item:

"Addition and subtraction of terms in sequence"

$2y + 7z - 8y - 4w + y + 14w - z =$

The Ss could not perform the step $3x + x = 4x$, nor indicate the sum or difference of unlike terms, $4x + 7y - y + 4x + 6y$. Competency one required that the S recognize that x is the same as 1x. For competency three test item, the S recognized equivalent terms. For example, that the term 6x can be combined with a like term 12x, but cannot be combined with an unlike term 5y. However, neither competency one nor three called for instruction in which the S actually combined equivalent terms, or showed the sum or difference of terms which are not equivalent. To remedy this, it was assumed that the existing definition of competency 6 required instruction in which the S would combine like terms $3x + x = 4x$, and show the sum or difference of unlike terms, $3x + 7y + x - y = 4x + 6y$.

Competency 10 description and test item:

"Simplifying an equation by adding and subtracting terms to both sides"

Solve for b: $7b + 2a + 3b - a = 10b + 2a + 35 - b$

The difficulty encountered at competency level ten was much like that described for competency level six, above. Ss who participated in the individual tryouts of the program could not perform the steps (a) collect like terms,

$8g + 2g = (8 + 2)g$; or (b) show the sum or difference of unlike terms: $8g + 2g + 3h = (8 + 2)g + 3h$. The existing definition of competence level ten was assumed to include instruction in collecting terms, and the additional instruction at competency level six, provided for showing the sum or difference of unlike terms.

Competency 7 description test item:

$$2 + 7 - 5 + 1 - 6 = 3 - 6 + 8 + 2 \ (\pm\ ?)$$

The instruction required to teach for the lower competencies was not sufficient preparation for the teaching of competency seven. The Ss could not combine signed positive and signed negative numbers to balance the sides of this equation. Competency four required instruction in the rules for combining signed numbers in sequence, and the actual performance of combining signed numbers in sequence: $5 - 4 + 2 - 2 + 6 = +7$. Additional practice was provided at competency level four, and it was assumed that competency seven would require instruction in an intermittent step, in which the signed numbers were combined on each side of the equation, before the missing signed number was supplied, as in this example.

$$2 + 7 - 5 + 1 - 6 = 3 - 6 + 8 + 2 \ (\pm\ ?)$$
$$-1 = +7 \quad (\pm\ ?)$$

APPENDIX F

RETENTION TEST

SIMPLIFYING EQUATIONS

Date ______________________________ Name ______________________

1. Solve for r: $8r + 2 = 7r + 5$

2. Solve for a: $6a + 2b = 5a + 4b$

3. Solve for v: $9v + 2w = 15w + 8v - 3w$

Form RT

Name ________________

4. Solve for x: $5x + 4y = 3x + 9y + x$

5. Solve for m: $7m - 3n = 9n - 5m - 8n + 11m$

6. Solve for e: $10e + 2f - 6e = 8f + 3e + 2f$

7. Solve for g: $2g + 4h + 5h = 12h + g$

Form RT

Name ____________________

8. Solve for c: $8c + 3d - c = 10d + 6c + 14$

9. Solve for s: $4s + 2t + 3s - 3t = 6s + 9t + 25 - t$

10. Solve for y: $6y + 4z + 3y - z = 9y + 4z + 40 - y$

APPENDIX G

LEARNING PROGRAM QUESTIONNAIRE

LEARNING PROGRAM QUESTIONNAIRE

Name ______________________

Date ______________________

What are your feelings about the learning program you have just finished? Please complete this questionnaire by checking in one of the boxes to the right of each question.

Do not check the boxes according to your overall impression. Instead, consider each question separately from the others.

Answer truthfully, and from your own point of view. The results of this questionnaire will not affect how well you did on the learning program.

	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
1. The learning program was interesting.					
2. The learning program was an enjoyable way to learn.					
3. The learning program was difficult.					
4. The learning program was understandable.					
5. The learning program was a "good teacher".					

APPENDIX H

SCORES FROM SUCCESSIVE TRYOUTS OF THE PROGRAM

SCORES FROM INDIVIDUAL TRYOUTS OF THE LEARNING PROGRAM WITH THREE STUDENTS OF HIGH, AVERAGE, AND LOW MENTAL ABILITY (N=3)

Ability Level	I.Q.	Time-to-Completion (hrs/min)	Competency Test Score
High	117	4:15	7
Average	103	5:10	5
Low	86	6:00	4
Mean ($\bar{X}$)	102.00	5:08	5.33

Note.--Since the purpose of individual tryouts of the program is to uncover program inadequacies by having the student answer aloud and question ambiguous portions of the program, no record of program errors is kept nor any posttest administered.

SCORES FROM SMALL GROUP ADMINISTRATION OF THE LEARNING PROGRAM TO THREE STUDENTS OF HIGH, AVERAGE, AND LOW MENTAL ABILITY (N=3)

Ability Level	I.Q.	Time-to-Completion (hrs/min)	Program Errors	Competency Test Score	Posttest Score
High	121	3:05	37	9	8
Average	101	4:05	61	5	8
Low	83	3:20	65	5	1
Mean ($\bar{X}$)	101.67	3:30	54.33	6.33	5.67

SCORES FROM THE SUMMATIVE EVALUATION
OF THE LEARNING PROGRAM
(N=9)

	I.Q.	Time-to-Completion (hrs/min)	Program Errors	Competency Test Score	Posttest Score
	111	4:30	11	10	10
	109	3:10	10	9	10
	109	3:10	14	9	6
	99	2:40	22	9	1
	97	4:35	47	5	0
	95	2:30	23	9	6
	91	3:40	72	7	6
	82	3:00	52	6	0
	80	2:45	32	7	5
Mean ($\bar{X}$)	97.00	3:20	31.44	7.89	4.89

APPENDIX I

RAW SCORES FOR I.Q. AND FOR THE SIX DEPENDENT VARIABLES BY SEQUENCE GROUP AND MENTAL ABILITY LEVEL

I.Q. SCORES BY SEQUENCE GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group		
	Forward	Reverse	Random
High	(N=7)	(N=6)	(N=6)
	126	123	125
	121	120	121
	120	116	116
	116	112	114
	111	110	111
	109	105	106
	105		
Average	(N=7)	(N=7)	(N=7)
	104	104	104
	101	102	103
	98	98	100
	95	96	98
	94	94	94
	93	93	93
	92	92	92
Low	(N=6)	(N=7)	(N=7)
	90	90	91
	89	90	90
	89	89	89
	86	86	87
	82	83	85
	81	81	81
		80	80

TIME TO COMPLETE THE INSTRUCTION (MINUTES) RAW SCORES BY SEQUENCE GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group		
	Forward	Reverse	Random
	(N=7)	(N=6)	(N=6)
	150	195	240
	125	180	170
	280	260	160
High	195	165	190
	145	210	240
	195	105	305
	155		
	(N=7)	(N=7)	(N=7)
	240	195	280
	205	220	170
Average	155	175	175
	160	205	210
	135	320	165
	260	240	220
	200	185	205
	(N=6)	(N=7)	(N=7)
	165	315	190
	220	365	145
	210	190	195
Low	155	195	270
	270	195	435
	250	205	210
		155	135

PROGRAM ERROR RAW SCORES BY SEQUENCE
GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group		
	Forward	Reverse	Random
	(N=7)	(N=6)	(N=6)
	6	6	28
	4	7	23
	16	20	16
High	4	23	25
	0	0	27
	55	59	20
	20		
	(N=7)	(N=7)	(N=7)
	23	42	38
	49	50	29
	10	6	54
Average	4	5	19
	50	63	48
	29	19	49
	34	29	49
	(N=6)	(N=7)	(N=7)
	17	114	13
	15	13	57
	7	30	47
Low	72	63	72
	20	23	48
	42	19	92
		60	140

COMPETENCY TEST RAW SCORES BY SEQUENCE GROUP
AND ABILITY LEVEL

Mental Ability Level	Sequence Group: Forward	Reverse	Random
	(N=7)	(N=6)	(N=6)
	9	10	8
	9	10	6
	10	7	8
High	10	8	9
	9	10	10
	8	7	10
	9		
	(N=7)	(N=7)	(N=7)
	8	8	9
	5	8	9
	10	8	7
Average	10	10	8
	6	5	6
	8	9	7
	7	6	8
	(N=6)	(N=7)	(N=7)
	9	4	10
	10	10	9
	8	9	8
Low	7	5	4
	9	8	7
	6	9	7
		4	4

ATTITUDE MEASURE RAW SCORES BY SEQUENCE GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group		
	Forward	Reverse	Random
High	(N=7)	(N=6)	(N=6)
	70	90	70
	85	80	70
	95	80	75
	80	80	75
	85	85	55
	55	45	80
	70		
Average	(N=7)	(N=7)	(N=7)
	60	45	50
	70	45	65
	55	80	55
	65	80	65
	40	65	40
	55	50	60
	65	55	70
Low	(N=6)	(N=7)	(N=7)
	60	20	80
	65	70	50
	80	30	85
	60	65	85
	65	90	90
	75	65	60
		90	45

POSTTEST RAW SCORES BY SEQUENCE
GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group: Forward	Reverse	Random
	(N=7)	(N=6)	(N=6)
	10	9	5
	1	8	7
	3	6	8
High	10	10	6
	5	5	9
	5	0	3
	4		
	(N=7)	(N=7)	(N=7)
	3	0	3
	5	2	4
	4	4	2
Average	6	6	3
	5	3	7
	3	4	4
	4	7	3
	(N=6)	(N=7)	(N=7)
	9	1	9
	5	6	0
	9	4	1
Low	1	5	3
	4	8	0
	4	0	3
		2	2

RETENTION TEST RAW SCORES BY SEQUENCE GROUP AND ABILITY LEVEL

Mental Ability Level	Sequence Group		
	Forward	Reverse	Random
High	(N=7)	(N=6)	(N=6)
	10	10	5
	8	9	5
	3	4	9
	10	10	10
	3	9	8
	0	0	4
	4		
Average	(N=7)	(N=7)	(N=7)
	0	2	3
	3	0	7
	3	6	2
	0	10	4
	4	2	4
	3	0	0
	0	6	10
Low	(N=6)	(N=7)	(N=7)
	10	0	10
	6	0	1
	10	6	4
	0	2	4
	4	9	0
	7	0	3
		2	1

REFERENCES

REFERENCES

Bloom, B. S. (Ed.) Taxonomy of educational objectives. Handbook I: cognitive domain. New York: David McKay, Inc., 1956.

Briggs, L. J. Sequencing of instruction in relation to hierarchies of competence. Pittsburgh, Pa: American Institutes for Research, 1968. Monograph No. 3.

Briggs, L. J. Handbook of procedures for the design of instruction. Pittsburgh, Pa: American Institutes for Research, 1970. Monograph No. 4.

Coleman, L. T. A study of the acquisition of the intellectual skill comparing in the area of social sciences. Unpublished Ph.D. dissertation, University of California, Berkeley, 1969.

Esbensen, T. Working with individualized instruction. The Duluth experience. Palo Alto, Cal: Fearon Publishers, 1968.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M., & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs: General and Applied, 1961, 75. (Whole No. 518).

Gagné, R. M., and Staff, University of Maryland Mathematics Project. Some factors in learning nonmetric geometry. Monographs of the Society for Research in Child Development, 1965, 30, 42-49.

Gagné, R. M., Mayor, J. R., Garstens, H. L., & Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs: General and Applied, 1962, 76. (Whole No. 526).

Hackett, M. A hierarchy of skills in listening comprehension and reading comprehension. Unpublished Ph.D. dissertation, University of California, Berkeley, 1968.

Levin, G. R., & Baker, B. L. Item scrambling in a self-instructional program. Journal of Educational Psychology, 1965, 54, 138-143.

Mager, R. F. On sequencing of instructional content. Psychological Reports, 1961, 9, 405-413.

Mager, R. F., & Clark, C. Explorations in student-controlled instruction. Psychological Reports, 1963, 13, 71-76.

Niedermeyer, F. C. The relevance of frame sequence in programmed instruction: an addition to the dialogue. AV Communication Review, 1968, 16, 301-317.

Niedermeyer, F. C., Brown, J., & Sulzen, B. Learning and varying sequences of ninth-grade mathematics materials. Journal of Experimental Education, 1969, 37, 61-66.

Payne, D. A., Krathwohl, D. R., & Gordon, J. The effect of sequence on programmed instruction. American Educational Research Journal, 1967, 4, 125-132.

Pressey, S. L. A simple apparatus which gives tests and scores--and teaches. School and Society, 1926, 13, 373-376.

Resnick, L. B., & Wang, M. C. Approaches to the validation of learning hierarchies. Paper presented at the Eighteenth Annual Western Regional Conference on Testing Problems, San Francisco, May, 1969.

Roe, K. V., Case, H. W., & Roe, A. A scrambled versus ordered sequence in autoinstructional programs. Journal of Educational Psychology, 1962, 53, 101-104.

Skinner, B. F. Teaching machines. Science, 1958, 128, 137-158.

Spady, W. G., & Greenwood, D. Instant path analysis: An elementary cookbook. Unpublished manuscript, The Ontario Institute for Studies in Education, 1969.

Wodtke, K., Brown, B., Sands, H., & Fredericks, P. The effects of subject-matter and individual difference variables on learning from scrambled versus ordered instructional programs. Paper presented at the meeting of the American Educational Research Association, Chicago, 1968.

VITA

Rosemary Yarr Spencer was born on June 22, 1937, in Chicago, Illinois. In 1959 she received her Bachelor of Science degree from The Florida State University with a major in elementary education. In 1960 she was awarded the degree of Master of Science in elementary education from the same university. Her doctoral program in the Department of Educational Research at The Florida State University was in the area of Instructional Systems with emphasis on research, development and evaluation for promoting effective learning.

For six years she taught in the public schools of Leon County, Florida, in the field of elementary education, grades one through six. In 1968 she returned to the doctoral program at The Florida State University. As part of her doctoral training, she had a year of supervised research, and another year in charge of all nonacademic testing at The Florida State University. Also, she developed and tested a programmed instruction text for the Elementary Education Model at The Florida State University.

She is a member of the Consultant Committee of the Clara Lewis Scholarship House for Girls at The Florida State University, Leon County Classroom Teachers Association, Florida Educational Research Association, and American Educational Research Association.

Mrs. Spencer is married to Platt Rogers Spencer, a Certified Public Accountant.